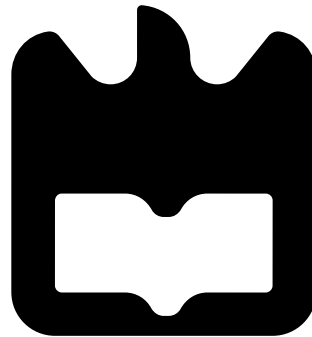




**João
Davim**

Ethernet para a Rede de Transporte





João
Davim

Ethernet para a Rede de Transporte

“Ever more startled at José Arcadio Buendía’s lucidity, he asked him how it was possible that they had him tied to a tree.

-Hoc est simplicissimus,- he replied - Because I’m crazy”

— Gabriel García Marquez



**João
Davim**

Ethernet para a Rede de Transporte

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Engenharia de Computadores e Telemática, realizada sob a orientação científica de Armando Pinto, Professor do Departamento de Electrónica Telecomunicações e Informática da Universidade de Aveiro

o júri / the jury

presidente / president

Prof. Doutor João Nuno Pimentel da Silva Matos

Professor Associado do Departamento de Electrónica, Telecomunicações e Informática da Universidade de Aveiro

vogais / examiners committee

Doutor Eng. Jorge Castro

Investigador INESC Porto

Prof. Doutor Armando Humberto Moreira Nolasco Pinto

Professor Auxiliar do Departamento de Electrónica, Telecomunicações e Informática da Universidade de Aveiro

**agradecimentos /
acknowledgements**

Queria deixar aqui um agradecimento para os meus pais, que sempre me apoiaram ao longo de todo o meu (longo) percurso escolar. Tenho de agradecer também à minha irmã, pela paciência e tempo que sempre teve para perder comigo.

Quanto aos meus colegas, passámos muito tempo juntos, muitas vezes a trabalhar, outras nem por isso, não tenho espaço para nomear e agradecer a toda a gente, fica aqui só um agradecimento ao Pedro Alves que acompanhou este trabalho, e o meu percurso universitário, mais de perto. De qualquer maneira, obrigado a vocês todos, estes 6 anos tinham sido muito aborrecidos, e eu seria uma pessoa muito mais pobre senão fossem vocês.

Queria também agradecer ao professor Armando Pinto que aceitou orientar este trabalho

Resumo

Actualmente, as redes de transporte utilizam principalmente tecnologia SDH sobre uma camada WDM. Esta dissertação aborda a mudança da tecnologia das redes de transporte, de SDH para Carrier Ethernet. Depois de descrever as limitações que a tecnologia Ethernet tem ao operar em redes de transporte, passamos a descrever a definição de tecnologia Carrier Ethernet segundo o Metro Ethernet Forum. De seguida são apresentadas dois protocolos propostos para suportar a tecnologia Carrier Ethernet.

Na segunda parte da dissertação é apresentado um modelo IPTV que combina o envio de canais em multicast e unicast para melhor utilizar os recursos da rede, de seguida o modelo é testado e é analisado o desempenho considerando o envio em unicast, em multicast e usando um modelo misto.

Abstract

Nowadays, most of the transport networks are based on SDH technology over a WDM layer. This dissertation approaches the undergoing change in transport networks from SDH to Carrier Ethernet. We describe the limitations that the Ethernet technology faces on a transport level. We proceed by describing the concept of Carrier Ethernet according to the Metro Ethernet Forum. Afterwards two protocols intended to support the Carrier Ethernet technology are presented.

On the second part of this dissertation we present a model that combines the use of multicast and unicast to use the resources of the network more effectively. After the model is presented then we test it, and analyze the performance of sending the channels using unicast, multicast and a mixed model.

Conteúdo

Conteúdo	i
Lista de Figuras	iii
Lista de Tabelas	v
1 Introdução	1
1.1 Objectivos	2
1.2 Estrutura da dissertação	3
2 Carrier Ethernet	4
2.1 Definição de Carrier Ethernet	6
2.2 Conceitos introduzidos em Carrier-Ethernet	7
2.2.1 <i>User-Network Interface</i> , UNI	7
2.2.2 <i>Ethernet Virtual Connection</i> , EVC	8
2.3 Atributos relativos a requisitos de Rede	9
2.3.1 Serviços Normalizados	9
2.3.2 Qualidade de Serviço	12
2.3.3 Gestão de Serviços	12
2.4 Atributos relativos a requisitos de Serviços	13
2.4.1 Escalabilidade	13
2.4.2 Recuperação e Protecção de Falhas	13
2.5 Tecnologias Ethernet para a área de transporte	14
2.6 Solução proposta pelo IEEE	16
2.6.1 IEEE 802.1AD - Provider Bridges	16
2.6.2 IEEE 802.1AH - Provider Backbone Bridges	17
2.6.3 IEEE 802.1Qay - Provider Backbone Bridges - Traffic Engineered	19
2.7 Solução proposta em conjunto por IETF/ITU-T	21
2.7.1 Diferenças entre MPLS e MPLS-TP	21
2.7.2 Aplicação de MPLS-TP a Carrier Ethernet	22
3 Modelo eficiente de distribuição de IPTV	24
3.1 Modelo de distribuição de IPTV	25
3.2 Modelar Sistema de IPTV	25
3.3 Cenário dinâmico e cenário estático	26
3.4 Abordagem Exacta	26
3.4.1 Cenário Estático	27

3.4.2	Cenário Dinâmico	29
3.5	Aproximação usando a distribuição Normal	31
3.5.1	Cenário Estático	33
3.5.2	Cenário Dinâmico	34
3.6	Optimização dos parâmetros do Modelo	36
3.6.1	Parâmetros do modelo	37
4	Resultados	38
4.1	Cenário Estático	38
4.1.1	Calcular o valor de M	38
4.1.2	Influência do valor de α	39
4.1.3	Influência do número de utilizadores activos	39
4.2	Cenário Dinâmico	40
4.2.1	Variação de β	41
4.2.2	Variação de α	42
5	Conclusões	45
5.1	Cenário Estático	46
5.2	Cenário Dinâmico	46
5.3	Trabalho Futuro	47
A	Ethernet	48
A.1	Elementos da Rede	48
A.1.1	Ethernet Switches e Bridges	49
A.2	Ethernet <i>frames</i>	49
A.3	Funcionamento Ethernet	50
A.4	IEEE 802.1Q - Virtual LAN	51
B	SDH	53
B.1	SDH	53
B.1.1	Elementos da Rede	54
B.1.2	SDH <i>frames</i>	54
B.1.3	Funcionamento SDH	54
C	C - Métodos probabilísticos	57
C.1	Variável de Bernoulli	57
C.2	Distribuição Binomial	57
C.3	Distribuição Multinomial	57
C.4	Distribuição de Zipf	58
C.4.1	Implicação do parâmetro β	58
D	Multi-Protocol Label Switching - MPLS	60
D.0.2	Equipamentos	61
D.1	Estabelecimento de Ligações e <i>multicast</i>	63
	Bibliografia	65

Lista de Figuras

2.1	Interligação entre as normas do MEF e a arquitectura da rede	7
2.2	Esquema interno de uma UNI	8
2.3	Exemplo de uma UNI que separa a rede de um cliente da rede do operador. .	8
2.4	Exemplo de E-Line.	9
2.5	Exemplo de E-LAN.	10
2.6	Exemplo de E-TREE.	10
2.7	Topologia verdadeira, e topologia processada pelo STP de uma rede de transporte conjuntamente com redes locais.	15
2.8	Evolução das tecnologias para a área de transporte	16
2.9	<i>Frame</i> usada pelo protocolo 802.1AD, Provider Bridges	16
2.10	Aplicação da tecnologia Provider Bridges.	17
2.11	Provider Backbone Bridge <i>frame</i>	17
2.12	Esquema de uma Backbone Edge Bridge. Imagem disponível em [6]	19
2.13	Relação entre o espaço PBB-TE e a entidade de gestão.	20
2.14	Cabeçalhos usados pelos protocolo MPLS e MPLS-TP	21
3.1	Camadas do sistema de IPTV	25
3.2	Função de distribuição de probabilidade de $Pr[n_u = i]$, isto é, a probabilidade de haver i canais enviados em unicast	28
3.3	Funções de distribuição de probabilidade de $Pr[n_u = j]$ e $Pr[n_m = i]$, ou seja a probabilidade de haverem j canais unicast e i canais multicast	31
3.4	Variação da quantidade de vectores B. Nota: Nos gráficos é usada uma escala logarítmica.	31
3.5	Comparação entre a abordagem exacta e a aproximação usando a distribuição normal.	32
3.6	Função de distribuição de probabilidade de $Pr[n_u = i]$, isto é, a probabilidade de haver i canais enviados em unicast	34
3.7	Funções de distribuição de probabilidade de $Pr[n_u = j]$ e $Pr[n_m = i]$, ou seja a probabilidade de haverem j canais unicast e i canais multicast	36
4.1	Evolução da carga no sistema, para diferentes distribuições de popularidade de canais	40
4.2	Evolução da carga no sistema dependendo do número de utilizadores activos .	42
4.3	Evolução da carga no sistema dependendo do valor de β	43
4.4	Evolução da carga no sistema dependendo do valor de α	44

5.1	Comparação entre o modelo misto unicast/multicast e uso só de unicast ou multicast	46
5.2	Comparação entre o modelo misto unicast/multicast e uso só de unicast ou multicast, para vários valores de utilizadores activos	47
A.1	camadas OSI	49
A.2	<i>frame</i> Ethernet	49
A.3	Frame VLAN	52
B.1	STS distribuido por duas <i>frames</i> SONET, imagem disponível em [21]	55
B.2	Exemplo de anel UPSR, imagem disponível em [21]	55
B.3	Exemplo de anel BLSR/4, imagem disponível em [21]	55
C.1	Vectores de distribuição de popularidade para diferentes valores de α	59
D.1	Enquadramento do protocolo MPLS no modelo OSI	60
D.2	Cabeçalho usado pelo protocolo MPLS	60
D.3	Pacote MPLS com cabeçalho e <i>payload</i>	61
D.4	Exemplo de uma rede que usa o protocolo MPLS, com 2 LSPs definidos	62

Lista de Tabelas

2.1	Ritmos de transmissão PDH.	4
3.1	Parâmetros do modelo [3]	37
4.1	Parâmetros dos cenários para iterativamente calcular o calor de M	39
4.2	Valores de M óptimos para diferentes distribuições de popularidade	41
4.3	Parâmetros dos cenários para iterativamente calcular o calor de M	41
4.4	Valores M óptimos para diferentes números de utilizadores activos	43
4.5	Parâmetros de diferentes cenários para calcular o impacto de β na carga exercida sobre a rede	43
4.6	Parâmetros dos cenários para iterativamente calcular o calor de M	44
A.1	Campos de uma <i>frame</i> Ethernet	50
B.1	Ritmos de transmissão SDH e SONET	54

Acrónimo	Descrição
ALNP	Aggregate Link and Node Protection
BCB	Backbone Core Bridge
BEB	Backbone Edge Bridge
B-BEB	Backbone Edge Bridge apenas com componente B
BSI	Backbone Service Instance
CBP	Costumer Backbone Port
CoS	Class of Service
EEPP	End-to-end Path Protection
ESP	Ethernet Switched Path
I-BEB	Backbone Edge Bridge apenas com componente I
LER	Label Edge Router
LSR	Label Switching Router
LLC	Logical Link Control
MAN	Metropolitan Area Network
MEN	Metro Ethernet Network
MP2MP	Multipoint-to-Multipoint
MPLS	Multiprotocol Layer Switching
MPLS-TP	Multiprotocol Layer Switching-Transport Profile
OAM	Operations, Administration and Maintenance
PB	Provider Bridges
PBN	Provider Bridge Network
PBB	Provider Backbone Bridges
PBB-TE	Provider Backbone Bridges - Traffic Engineered
PBBN	Provider Backbone Bridge Network
PEB	Provider Edge Bridge
QoS	Quality of Service
STP	Spanning Tree Protocol
T-MPLS	Transport-MPLS
UNI	User Network Interface
WAN	Wide Area Network

Capítulo 1

Introdução

As redes de transporte hoje em dia usam principalmente a tecnologia SDH sobre canais WDM. Esta tecnologia foi desenvolvida com o propósito de transportar sinais telefónicos substituindo as redes de cobre, mas permitindo velocidades de transmissão mais elevadas.

No entanto com a evolução de serviços há, actualmente necessidade das redes de transporte suportarem de forma eficiente tráfego sob forma de comunicações do tipo ponto-multiponto e serem mais flexíveis em termos de largura de banda oferecida. A tecnologia SDH não foi projectada para operar neste cenário, e a tecnologia Ethernet que consegue preencher estes requisitos não consegue operar satisfatoriamente em redes de transporte.

É uma situação de compromisso, em que é preciso abdicar de um pouco da simplicidade da tecnologia Ethernet para ganhar capacidades de rede de transporte, ou alterar o SDH para ser mais flexível. Assim, que vantagens oferece a Ethernet sobre as restantes soluções?

- Ethernet é a tecnologia dominante no mercado LAN, na realidade, mais de 90% das LAN usam Ethernet [16] e segundo o Metro Ethernet Forum 99% do tráfego que circula nas redes MAN e WAN tem origem e destino numa porta Ethernet [16]. Esta situação coloca o problema de transportar *frames* ou células com formato e tamanho diferentes das *frames* originais. Actualmente sempre que uma *frame* é encaminhada para fora da LAN, ou é convertida no formato adequado e sempre que a tecnologia de encaminhamento muda é reconvertida, e depois uma eventual resposta passa por um processo idêntico, com a desvantagem de este processo de conversão aumentar o atraso sofrido pelas *frames* que atravessam as redes de transporte[15]. Outra solução é o encapsulamento sucessivo das *frames* de acordo com a tecnologia das redes onde passam, o problema desta solução é o *overhead* introduzido na rede com os cabeçalhos adicionados. Com uma tecnologia comum na camada de ligação de dados, estes problemas desaparecem porque o formato das *frames* é o mesmo quer nas redes locais, quer nas redes de transporte. Assim não há os atrasos introduzidos pela conversão das *frames* entre tecnologias, nem desperdício de largura de banda com pilhas de cabeçalhos [16]
- Ethernet é um mercado em grande escala onde a concorrência faz com que haja preços mais baixos. A tecnologia Ethernet é um *open-standard* em que todos os aspectos estão normalizados e qualquer fabricante pode produzir equipamentos pagando apenas uma pequena quantia ao IEEE. Esta política de *open-standard* conduziu a uma total interoperabilidade dos equipamento, e a um baixar dos custos de *hardware*. Estender

esta lógica ao mercado das redes de transporte, trará redução de custos com o equipamento, e o fim de problemas com a interoperabilidade entre equipamentos de diferentes fabricantes.

- Outro aspecto onde a tecnologia Ethernet tem vantagem sobre a tecnologia SDH enquanto solução para as redes de transporte é na formação da mão-de-obra, pois apenas é preciso formação numa tecnologia, neste caso Ethernet, e não em duas como actualmente. [16],

Mas a tecnologia Ethernet não opera satisfatoriamente nas redes de transporte, a tecnologia Carrier Ethernet surge assim como uma solução para lidar com as novas exigências a que as redes de transporte têm de responder. O Metro Ethernet Forum define Carrier Ethernet como um serviço normalizado, ubíquo de transporte e de rede, definido por 5 atributos que distinguem Carrier Ethernet da Ethernet tradicional:

- Serviços Normalizados
- Escalabilidade
- Recuperação e Protecção de Falhas
- Qualidade de Serviço
- Gestão de Serviços

Estes atributos permitem à tecnologia Carrier Ethernet oferecer funcionalidades de gestão, recuperação de falhas, e escalabilidade que a Ethernet tradicional não tem, e que associados à flexibilidade e simplicidade da Ethernet tradicional permitem à tecnologia Carrier Ethernet assumir-se como alternativa à tecnologia SDH para as redes de transporte.

Até ao momento o MEF lançou mais 25 especificações que contêm especificações técnicas sobre os mais diversos temas como OAM, serviços e atributos de serviços, ou *swites* de teste.

Apesar do MEF ser a organização mais empenhada em normalizar e divulgar a tecnologia Carrier Ethernet, não está sozinho encontrando-se a trabalhar em conjunto com o IEEE e com o ITU-T para o desenvolvimento de normas de forma cooperativa, de modo a garantir uma grande aceitação das mesmas. Entre as empresas associadas ao MEF, estão principalmente fornecedores de serviços de telefone, Internet e TV, e fabricantes de equipamentos. Ambos os tipos de empresas vêm vantagem em associar-se ao MEF e certificar os seus serviços/produtos como *Carrier Ethernet Compliant*, porque a certificação garante a compatibilidade de serviços e interoperabilidade do equipamento.

1.1 Objectivos

Os objectivos desta dissertação são:

- Identificar as motivações para levar a Ethernet para a área do transporte
- Identificar as características e serviços prestados por uma rede Carrier Ethernet
- Identificar os pontos fortes e fracos da tecnologia Ethernet para a área do transporte

- Analisar de forma crítica o principal trabalho que tem vindo a ser desenvolvido pelos diversos organismos de normalização a actuar nesta área
- Analisar formas eficientes de disponibilizar serviços de IPTV usando Carrier Ethernet na rede de transporte

1.2 Estrutura da dissertação

A estrutura da dissertação é a seguinte: no capítulo seguinte, é definido o conceito de Carrier Ethernet assim como os problemas de usar Ethernet nas redes de transporte e as soluções propostas. No terceiro capítulo é apresentado um modelo para distribuir sinais de IPTV usando uma combinação de unicast e multicast. No quarto capítulo são apresentados os resultados e no quinto e último capítulo são apresentadas as conclusões.

Capítulo 2

Carrier Ethernet

As redes de transporte actuais usam principalmente a tecnologia SDH. Esta tecnologia foi desenvolvida no início dos anos 1980, e começou a ser normalizada em 1985 [17]. Na altura a quase totalidade do tráfego provinha de ligações telefónicas, que eram codificadas a uma taxa de 64 kbit/s usando a tecnologia PDH. Este sinal de 64 kbps era obtido dividindo um sinal de 2.048 Mbps por 32 canais, o sinal de 2.048 Mbps tem a designação de E1. Na altura a tecnologia PDH foi desenvolvida em conjunto com a instalação de fibra óptica nas redes de transporte porque a fibra óptica permite velocidades de transmissão muito mais elevadas do que as permitidas usando ligações de cobre. A tecnologia SDH era usada para transmitir os sinais PDH depois de estes serem agregados em sinais de ritmos superiores. A tecnologia PDH também suporta ritmos mais elevados e que são múltiplos do ritmo de transmissão E1, na tabela 2.1 estão os ritmos de transmissão da tecnologia PDH. Para manter o sincronismo entre os sinais, quando se agregam os sinais de mais baixo ritmo em sinais de ritmo mais elevado são introduzidos bits extra, esta operação é conhecida por *bit stuffing*.

Tabela 2.1: Ritmos de transmissão PDH.

Sinal PDH	Velocidade (Mbps)
E1	2
E2 (4 x E1) + <i>bit stuffing</i>	8
E3 (4 x E2) + <i>bit stuffing</i>	34
E4 (4 x E3) + <i>bit stuffing</i>	140
E5 (4 x E4) + <i>bit stuffing</i>	565

Eventualmente os ritmos de transmissão oferecidos pela tecnologia PDH tornaram-se insuficientes para transportar a quantidade de informação que circulava nas redes de transporte. Foi então introduzida a tecnologia SDH, que actuando sobre uma camada WDM, e usando os sinais PDH como tributários consegue alcançar débitos elevados.

Nos anos 90 com o aparecimento e expansão da Internet, o tráfego baseado em pacotes tornou-se comum, e agora é o tipo de tráfego mais comum nas redes de transporte. No entanto

a tecnologia SDH tem vários problemas em lidar com redes de comutação de pacotes, uns derivados de ter sido uma tecnologia pensada para lidar com sinais telefónicos, outros devido a ser uma tecnologia pouco flexível. De seguida estão enumeradas as desvantagens do uso da tecnologia SDH nas redes de transporte de futuro.

1. Pouca flexibilidade na largura de banda oferecida, que está limitada aos ritmos STS/STM indicados na tabela B.1 no Anexo B, assim um cliente que tem necessidade de largura de banda que não se enquadre num desses ritmos ou tem de pagar por um serviço que não está a usufruir, ou fica limitado a um serviço que não satisfaz as suas necessidades.
2. OPEX: os custos com os pessoal técnico constitui uma parte considerável dos custos de operação de uma rede, e as redes SDH requerem bastante configuração manual da rede o que implica técnicos treinados, e custos de deslocação (*truckrolls*) até à localização do equipamento[16].
3. SDH é orientado para velocidades de transmissão fixas, as *frames* são enviadas a cada $125\mu s$, e esse limite é imposto por hardware, mesmo por pequenos instantes é impossível ritmos de transmissão acima do ritmo STS/STM estipulado
4. SDH é também orientado para ligações extremo-a-extremo, não suportando originalmente tráfego do tipo ponto-a-multiponto, ou multiponto-a-multiponto, este facto é uma limitação para redes de transporte modernas em que a distribuição de TV e videoconferência são requisitos fundamentais.

A tecnologia Ethernet surgiu nos anos 70 nos laboratórios da Xerox, mas a partir de 1980 o IEEE passou a coordenar o esforço de normalização desta tecnologia através do projecto IEEE 802[17]. Durante os anos que se seguiram a tecnologia Ethernet suplantou as tecnologias concorrentes no mercado LAN, a tecnologia *Token Ring* usada pela IBM, e a tecnologia *Token Bus* apoiada pela General Motors, e tornou-se um onnipresente no mercado LAN, de tal maneira que neste momento mais de 90% de todo o tráfego gerado tem origem numa porta Ethernet[16].

Mas o que define a tecnologia Ethernet? Formalmente a tecnologia Ethernet é definida pela norma IEEE 802.3, que define três parâmetros das redes Ethernet:

1. O meio de transmissão, fibra, cobre (cabo co-axial, par entrançado ou outro) ou o ar.
2. As *frames* usadas para a comunicação entre os dispositivos.
3. O protocolo usado para a comunicação entre os dispositivos, este protocolo lida com a transmissão e recepção de *frames*, detecção de erros e colisões e ainda a sinalização da comunicação.

À partida aplicar uma tecnologia simples e orientada para transportar tráfego sob a forma de pacotes às redes de transporte poderia resolver os problemas que as redes SDH enfrentam, mas a tecnologia Ethernet actual não pode ser aplicada ao mercado das redes de transporte. Por ser uma tecnologia demasiado focada em redes locais tem aspectos que a penalizam quando aplicada no universo das redes de transporte. As principais limitações da tecnologia Ethernet são:

- Tempo de recuperação, a tecnologia SDH recupera de falhas em 50 ms, usando o protocolo Spanning Tree o tempo que uma rede Ethernet demora a recuperar de uma falha varia entre 2 e 120 s
- Número de dispositivos, o protocolo Spanning Tree obriga a que o raio da rede (o número máximo de dispositivos onde uma *frame* pode passar) seja menor ou igual a 7.
- Escalabilidade em número de utilizadores, a tecnologia Ethernet implica que qualquer dispositivo presente na rede conheça os endereços MAC de todos os outros, quando o número de nós começa a aumentar as pesquisas para fazer o encaminhamento das *frames* podem introduzir um atraso muito grande na rede.
- QoS, o protocolo Ethernet não tem suporte nativo para QoS, visto que o próximo salto de uma frame é apenas determinado com base no endereço destino ou VLAN-ID. Usar VLAN permite definir prioridades por *tag* VLAN mas isso é insuficiente porque não há maneira de fazer com que as *frames* sejam entregues de acordo com algum SLA, nem há sequer garantia que a *frame* seja entregue.
- Suporte a serviços TDM, tecnologia Ethernet não tem suporte nativo para emular circuitos TDM, no contexto MAN/WAN este facto pode ser um problema porque circuitos TDM são uma fonte importante de receitas para os prestadores de serviços.
- Os equipamentos Ethernet tradicionais quando recebem uma *frame* com endereço destino Multicast, ou com endereço desconhecido, tratam-na como se fosse uma *frame* com endereço destino Broadcast, enviando-a por todas as portas. Este comportamento faz com que seja desperdiçada largura de banda, e torna o inútil o envio de tráfego Multicast.

Mesmo assim, há a esperança que alterando um pouco o funcionamento da tecnologia Ethernet, estas limitações sejam ultrapassadas, e que a tecnologia Ethernet consiga superar a tecnologia SDH nas redes de transporte.

2.1 Definição de Carrier Ethernet

O MEF, entidade que define os conceitos relacionados com Carrier Ethernet, define o conceito de Carrier Ethernet como:

Rede e Serviços normalizados, universais de rede de transporte definidos pelos cinco atributos que distinguem a tecnologia Carrier Ethernet da tecnologia Ethernet.

O conceito de Carrier Ethernet é definido de forma dual, por um lado a parte de Serviços que diz respeito aos serviços prestados pelas redes Carrier Ethernet, por outro a parte de Rede que diz respeito ao comportamento da rede. Na figura 2.1 está a interligação entre as normas do MEF e a arquitectura da rede.

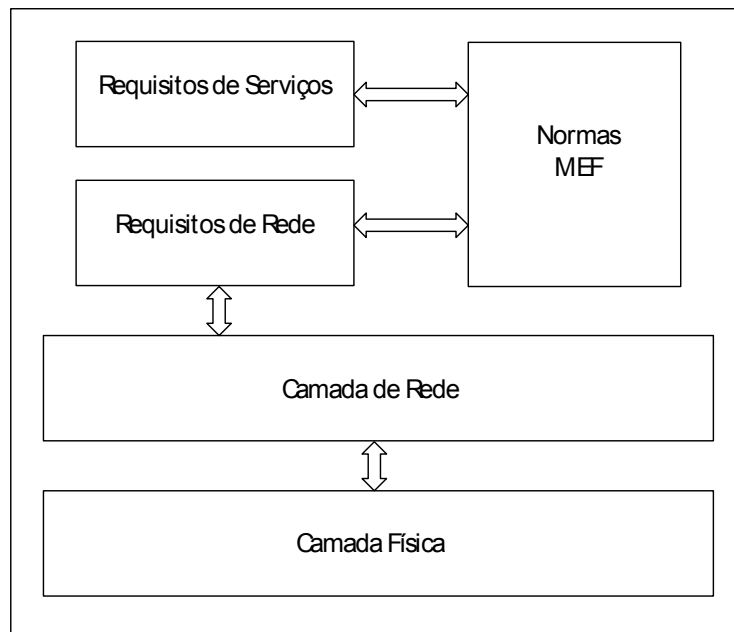


Figura 2.1: Interligação entre as normas do MEF e a arquitectura da rede

As normas e conceitos definidos pelo MEF são independentes das tecnologias de encaminhamento ou camadas de controlo da rede, apenas definem o que a rede deve fazer, ou os serviços que a rede deve prestar, **não** como o fazer. As normas aprovadas pelo MEF que estão numeradas cobrem areas como OAM, definição de serviços ou QoS. Os atributos que distinguem a tecnologia Carrier Ethernet da tecnologia Ethernet tradicional são:

- Serviços Normalizados
- Escalabilidade
- Recuperação e Protecção de Falhas
- Qualidade de Serviço
- Gestão de Serviços

2.2 Conceitos introduzidos em Carrier-Ethernet

O MEF introduziu dois novos conceitos na tecnologia Carrier Ethernet, que são a base da prestação de serviços entre o operador e o cliente. Estes conceitos são o: *User-Network Interface* (UNI) [16], e *Ethernet Virtual Connection* (EVC) [16].

2.2.1 *User-Network Interface*, UNI

Uma *User-Network Interface* (UNI) é a fronteira lógica onde a responsabilidade do prestador de serviços sobre a rede acaba e começa a do cliente. As UNI seguem o modelo com três camadas da figura 2.2.

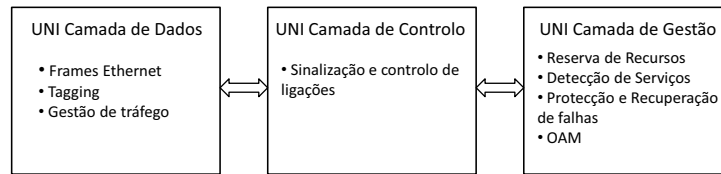


Figura 2.2: Esquema interno de uma UNI

Cada camada tem um conjunto de funções diferente:

- A Camada de Dados, é a responsável por implementar as camadas Física e Ligação de Dados (ver Anexo A).
- A Camada de Controlo, que define a forma como o cliente e o operador comunicam. Esta camada define os meios pelos quais os serviços são oferecidos ao cliente, isto resulta num acordo entre o cliente e o operador.
- A Camada de Gestão que configura e monitoriza a operação das camadas de Dados e de Gestão.

Há dois tipos de UNI:

- UNI-C (User Network Interface - Client side): Que contém todas as funções necessárias para ligar um cliente à rede.
- UNI-N (User Network Interface - Network side): Que contém todas as funções para ligar a rede aos clientes

2.2.2 *Ethernet Virtual Connection, EVC*

Um EVC é um construtor que guarda a informação de todas as UNIs que participam numa ligação, e impede a transferência de informação para UNIs que não façam estejam autorizadas a receber tráfego dessa ligação. O EVC guarda também a informação relativa ao SLA acordado para cada fluxo de dados.

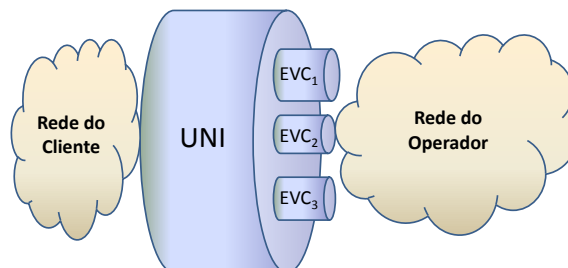


Figura 2.3: Exemplo de uma UNI que separa a rede de um cliente da rede do operador.

Cada UNI pode ter vários EVCs associados, como é visível na figura 2.3, cada um com SLAs diferentes a serem respeitados pelo prestador de serviços.

2.3 Atributos relativos a requisitos de Rede

2.3.1 Serviços Normalizados

O MEF definiu um atributo de Carrier Ethernet como serviços normalizados, estes serviços são um modo genérico de garantir conectividade na rede do operador e entre a rede do operador e as redes dos clientes:

- Circuitos virtuais que fornecem serviços E-Line, E-LAN e E-Tree de forma transparente Segundo as especificações do MEF a tecnologia Carrier Ethernet fornece três tipos de serviço virtuais, com estes três tipos de serviço é possível disponibilizar qualquer serviço modificando apenas os parâmetros da ligação. É ainda possível realizar multiplexagem de serviço, ou seja, o mesmo UNI estar ligado a mais que um serviço ao mesmo tempo, por exemplo um cliente pode estar ligado a um serviço E-Line que lhe proporciona acesso telefónico ao mesmo tempo que com o mesmo UNI está a participar numa videoconferência usando um serviço E-LAN.



Figura 2.4: Exemplo de E-Line.

- E-Line é um serviço de que usa uma EVC de ligação ponto-a-ponto para ligar dois UNIs através da rede e permite a comunicação entre eles, como mostrado na figura 2.4. Pode ser usado como uma Ethernet Private Line (EPL) ou Virtual Private Line (EVPL) para ligar dois UNIs através da rede ou pode servir para o prestador de serviços fornecer acesso Internet aos clientes. Apesar de um serviço E-LINE ligar apenas dois clientes, pode ser usado para transportar um grande número de serviços entre esses dois clientes, cada serviço tendo requisitos próprios.

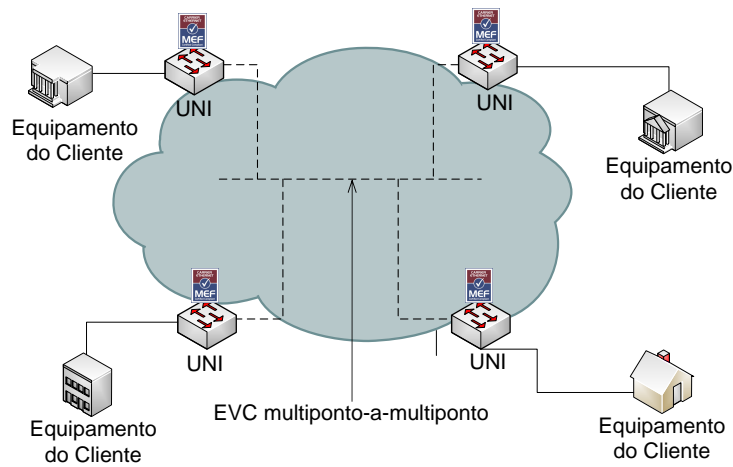


Figura 2.5: Exemplo de E-LAN.

- E-LAN é um serviço que permite comunicação multiponto-a-multiponto que serve para ligar vários UNIs através da rede. Permite comunicação entre todos os UNIs. Um exemplo pode ser visto na figura 2.5. Os perfis dos UNIs ligados por uma E-LAN podem ser bastantes díspares ao nível de largura de banda ou outros parâmetros de ligação, assim o EVC de ligação multiponto-a-multiponto pode guardar um perfil de QoS para cada UNI presente. Entre as utilizações possíveis para este serviço estão VPNs, serviço de LAN através da rede (permitir a comunicação de UNIs que estão na mesma LAN apesar de não estarem juntos fisicamente), e o fornecimento de serviços de Videoconferência ou outros que usem *multicast*.

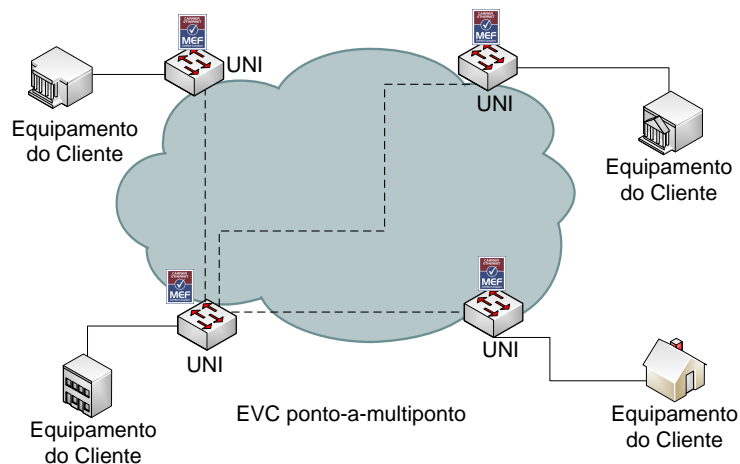


Figura 2.6: Exemplo de E-TREE.

- O último tipo de serviço aprovado pelo MEF é o E-TREE. A aprovação deste tipo

de circuito virtual foi adiada porque o E-TREE era considerado uma degeneração do serviço E-LAN. O serviço E-TREE tem uma UNI Raiz, e UNIs folha. É permitido o tráfego bidireccional entre as UNI folhas e a UNI raiz, mas não entre UNI folhas. Para oferecer redundância podem estar presentes mais que uma UNI raiz, neste caso, as UNI folha podem comunicar com as UNI Raiz, e as UNI Raiz podem comunicar entre si, mas a UNI folha continuam a não poder comunicar entre si. Um exemplo de uma E-TREE está na imagem 2.6. Entre as aplicações da E-TREE estão serviços de IPTV, acesso à internet e serviços *triple-play*.

- Oferta de serviços global e local através de equipamento normalizado Carrier Ethernet permite a prestação de serviços universais local e globalmente através de equipamento normalizado, ou seja os serviços fornecidos por Carrier Ethernet não devem depender de características específicas do *hardware* de nenhum fabricante, mas apenas dependentes de equipamento e serviços normalizados.

- Manutenção do equipamento LAN e redes do cliente sem problemas de conectividade Segundo o MEF, a adopção da tecnologia Carrier Ethernet por parte dos prestadores de serviços deve implicar a não necessidade de mudança de equipamento dos clientes, nem a problemas com conectividade nas redes locais dos clientes. Este requisito deve-se ao facto de se querer compatibilidade com a tecnologia já implementada, e minorar resistência à adopção da tecnologia. Caso fosse preciso mudar o equipamento terminal de todos os clientes, iria concerteza surgir resistência quer por parte dos operadores quer por parte dos clientes à adopção da tecnologia Carrier Ethernet.

- Granularidade de largura de banda, e de opções de qualidade de serviço

Para contrapor à pouca flexibilidade dos ritmos de transmissão da tecnologia SDH, a tecnologia Carrier Ethernet oferece largura de banda fléxivel com incrementos de 1M a partir de um mínimo de 1Mbps até um máximo de 10Gbps, isto é conseguido instalando *hardware* normalizado para 1G, e por *software* configura-se a largura de banda pretendida que pode, até atingir o limite imposto pelo *hardware*, ser alterado remotamente a pedido do cliente. A juntar a isto tem-se que é possível definir perfis de largura de banda como um tuplo do tipo [16]:

Perfil = < CIR, CBS, EBS, CM >

Em que:

- CIR (Committed Information Rate): velocidade média que as *frames* são entregues por tipo de performance
- CBS (Committed Burst Size): especifica a taxa média (em relação à CIR) de chegada de *frames* para a qual estas ainda são admitidas na rede. Estas *frames* são admitidas na rede sem qualquer garantia de performance.
- EBS (Excess Burst Size): é o limite, em bytes, para o número máximo de *frames* de serviço admitidas pela UNI
- CM (Color Marking): A *frames* de serviço são marcadas com cores diferentes de acordo com o critério que respeitam:

- * Verde: Respeita as taxas CIR e CBS
- * Amarelo: Respeita a taxa CBS e é entregue se a rede estiver pouco congestionada
- * Vermelho: A *frame* fora do perfil de largura de banda e é descartada

2.3.2 Qualidade de Serviço

As redes Ethernet tradicionais não tinham mecanismos de QoS, mas estes são essenciais em redes de transporte em que há possivelmente milhões de utilizadores, e cada um com SLAs que há que respeitar, assim a rede deve ter mecanismos para prestar os serviços acordados. Segundo o MEF o atributo de Qualidade de Serviço engloba 3 vertentes:

Granularidade de largura de banda, e de opções de qualidade de serviço

Este requisito sobrepõe-se nos atributos de Serviços Normalizados, e de Qualidade de Serviço. Na vertente de qualidade de serviço, este atributo é relevante no acordo entre o cliente e o prestador de serviços em relação à largura de banda disponibilizada porque a tecnologia Carrier Ethernet oferece largura de banda flexível com incrementos de 1M a partir de um mínimo de 1Mbps até um máximo de 10Gbps.

SLAs que fornecem serviços extremo-a-extremo

As redes Carrier Ethernet devem ter capacidade de fornecer serviços extremo-a-extremo com base no SLAs acordados com os clientes. Estes SLAs devem ser respeitados em termos de desempenho para a entrega de serviços como voz, dados e video, o restante tráfego é entregue em regime de *best-effort*. Podem também ser acordados SLAs relativos a parâmetros mais específicos da rede como CIR, número de *frames* perdidas, atraso e variação do atraso na entrega de *frames*.

2.3.3 Gestão de Serviços

Uma das maiores falhas das redes Ethernet era a falta de mecanismos de gestão de rede. O MEF consciente desse facto adicionou mecanismos de gestão à tecnologia Carrier Ethernet. Esta reflecte-se em 3 pontos:

- Capacidade de monitorizar, diagnosticar e centralmente gerir a rede usando protocolos independentes de fabricantes, isto garante interoperabilidade entre componentes, e permite abandonar as *truck-rolls* características das redes SDH, e permite gerir a rede remotamente
- OAM de classe de transporte, permite gerir ligações, e comunicações extremo-a-extremo
- Fornecimento rápido de serviços, permite a instanciação rápida e remota de serviços a um cliente quando este a solicita

2.4 Atributos relativos a requisitos de Serviços

2.4.1 Escalabilidade

Um dos requisitos da tecnologia Carrier Ethernet é a escalabilidade, ou seja, ter a capacidade de variar quer o número de utilizadores, quer o alcance geográfico, quer a largura de banda de um cliente sem degradar o desempenho geral da rede. A escalabilidade pode ser medida em 3 factores:

- Escalabilidade em termos de suportar milhões de utilizadores. Uma rede de transporte pode ter milhões de clientes, com várias CoS que têm de ser respeitadas, foram já desenvolvidas tecnologias que permitem ultrapassar a dificuldade das redes Ethernet tradicionais em lidar com grandes números de utilizadores como as tecnologias *Provider Bridges* e *Provider Backbone Bridges*.
- Escalabilidade em termos de alcance geográfico. A capacidade de fornecer serviços a áreas locais, metropolitanas e nacionais, ou seja capacidade de alcançar um grande espaço geográfico é essencial para uma rede de transporte para alcançar o maior número de clientes possível, permitindo assim a disponibilidade de serviços mais abrangentes.
- Escalabilidade em termos de largura de banda. A largura de banda disponível em Carrier Ethernet é oferecida de forma flexível com incrementos de 1Mbps a partir de um mínimo de 1Mbps até um máximo de 10Gbps. Sendo previsível que ligações a 10 Gbps não sejam disponibilizadas a clientes domésticos no futuro próximo, esta escalabilidade de largura de banda é importante para clientes industriais e para os segmentos *core* da rede onde o tráfego de ritmos mais baixos é transportado depois de agregado.

Estes três requisitos fazem a tecnologia Carrier Ethernet dar um salto em termos de capacidade da rede, e são essenciais para tornar a Carrier Ethernet uma tecnologia atractiva para os prestadores de serviços e para os clientes.

2.4.2 Recuperação e Protecção de Falhas

As redes ópticas SDH, ao contrário das redes Ethernet tradicionais, têm mecanismos de protecção recuperação bastante eficientes que garantem a protecção da comunicação e recuperação de falhas em menos de 50ms. Este valor tornou-se o tempo de referência para recuperação de falhas em redes de transporte, e um valor que a Carrier Ethernet tem de alcançar. O MEF define 3 requerimentos para a recuperação e protecção de falhas em redes Carrier Ethernet:

- Capacidade de detectar e recuperar de falhas sem que os utilizadores tenham conhecimento ou percepção da falha que ocorreu
- Responder aos requisitos de mais exigentes em termos de qualidade e disponibilidade
- Tempo de recuperação de falhas em menos de 50 ms

Em redes ópticas tradicionais a protecção era geralmente obtida com métodos de protecção 1+1 ou 1:1 que desperdiçavam metade da largura de banda instalada. Em Carrier Ethernet esses métodos de protecção foram abandonados, sendo substituídos por quatro novas formas de protecção, publicadas na norma MEF 2.

- Aggregate Link and Node Protection (ALNP) que oferece protecção contra falhas de nós e ligações locais
- End-to-end Path Protection (EEPP) que oferece protecção extremo-a-extremo onde são estabelecidos dois caminhos, um principal e um de protecção que é activado em caso de falha na rede.
- Multipoint-to-Multipoint (MP2MP) que fornece protecção para E-LAN
- Protecção de ligações por agregação em que ligações que liguem os mesmos nós podem ser agregadas

2.5 Tecnologias Ethernet para a área de transporte

Aplicar a tecnologia Ethernet às redes de transporte traz vantagens quer para operadores quer os clientes. Mas a tecnologia Ethernet, universal nas redes locais não pode ser aplicada directamente nas redes de transporte. Os problemas que impedem essa aplicação directa são:

- Escalabilidade: O problema da escalabilidade tem várias vertentes:
 1. Ethernet obriga a que os terminais guardem o endereço MAC de todos os terminais na rede, numa rede MAN/WAN com milhares de terminais, pesquisas e filtragens não podiam ser feitas em tempo útil pelos equipamentos
 2. Usar a tecnologia Ethernet nas redes MAN/WAN faz com que cada um dos 4094 VLAN IDs tenha de ser global na rede de transporte sob o risco de haver *frames* mal encaminhadas, e torna o número de VLAN IDs insuficiente.
 3. Quando uma Ethernet bridge recebe uma frame com um DA que não conhece envia-a por todas as portas, menos pela porta de onde a frame chegou. Numa rede com milhares de terminais a probabilidade de isto acontecer aumenta.
- Protecção: O encaminhamento numa rede Ethernet que use várias bridges usa o protocolo STP (ou uma das suas variantes: MSTP ou RSTP) e este tem um tempo de convergência que varia entre os 2 e os 120s, o que só por si o exclui como opção visto que o tempo de recuperação de referência é 50 ms. Mas ainda há a agravante de todas as bridges na rede participarem no protocolo, o que no caso da Ethernet de transporte inclui as bridges do fornecedor e as dos clientes porque o protocolo não faz essa distinção, o que tem como consequência a necessidade de todo protocolo reconvergir em todas as redes MAN/WAN e locais sempre que há uma mudança na topologia numa rede local. Um exemplo deste problema está na figura 2.7 em que uma rede de transporte que liga 2 redes locais, cada uma dividida em 2 localizações, e na parte de baixo da figura a topologia do ponto de vista do STP, sem fronteiras entre redes locais e de transporte.

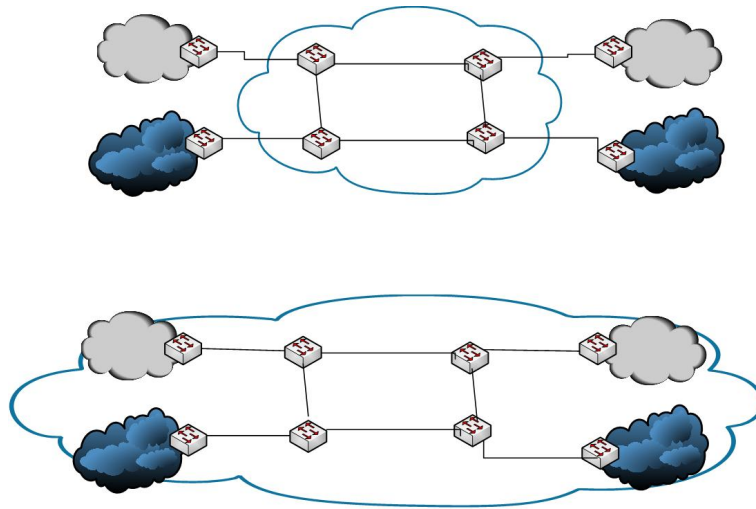


Figura 2.7: Topologia verdadeira, e topologia processada pelo STP de uma rede de transporte conjuntamente com redes locais.

- QoS: O protocolo Ethernet não tem suporte nativo para QoS, visto que o próximo salto de uma frame é apenas determinado com base no endereço destino ou VLAN-ID. Usar VLAN permite definir prioridades por VLAN tag mas isso é insuficiente porque não há maneira de fazer com que as *frames* sejam entregues de acordo com algum SLA, nem há sequer garantia que a *frame* entregue.

Assim, antes de poder aplicar a tecnologia Ethernet nas redes de transporte foi preciso evoluir a tecnologia Ethernet tradicional, esta evolução foi feita de forma concorrente por três organismos de normalização, IEEE, ITU-T e IETF. O IEEE optou por modificar a *frame* Ethernet para incluir *tags* que consigam resolver as limitações da tecnologia Ethernet. O ITU-T começou o também o desenvolvimento de uma tecnologia para a área de transporte, optando por alterar o protocolo MPLS, essa tecnologia designada por T-MPLS, foi abandonada após terem surgido problemas com a compatibilidade entre o T-MPLS e o MPLS. Para resolver esses problemas em 2008 o ITU-T juntou-se ao IETF para desenvolver em conjunto um protocolo que ficou conhecido como MPLS-Transport Profile, ou MPLS-TP. Na figura 2.8 está um diagrama para a evolução das tecnologias para a área de transporte.

De seguida vão ser descritas as tecnologias actuais para a rede de transporte. Como foi dito acima há neste momento duas abordagens, a do IEEE que passa por três normas diferentes até chegar a um protocolo final, e a solução conjunta do IETF e ITU-T, na lista abaixo estão os nomes das soluções propostas:

- Solução IEEE
 - IEEE 802.1AD - Provider Bridges
 - IEEE 802.1AH - Provider Backbone Bridges
 - IEEE 802.1Qay - Provider Backbone Bridges - Traffic Engineered
- Solução IETF/ITU-T
 - MPLS-TP

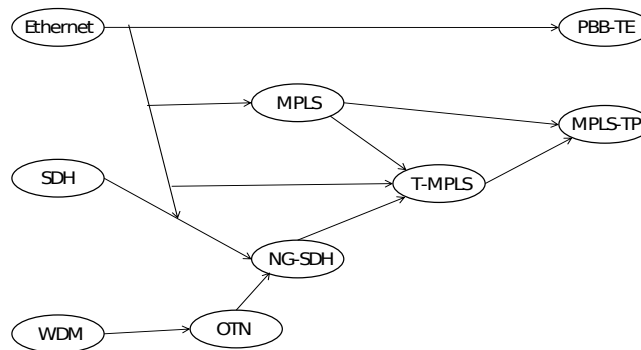


Figura 2.8: Evolução das tecnologias para a área de transporte

2.6 Solução proposta pelo IEEE

Nesta secção irão ser explicados os passos dados pelo IEEE até chegar a uma solução capaz de, usando a tecnologia Ethernet, ser aplicada com eficácia às redes de transporte. A solução adoptada pelo IEEE foi adicionar sucessivamente *tags* para resolver os problemas de aplicar a tecnologia Ethernet às redes de transporte.

2.6.1 IEEE 802.1AD - Provider Bridges

Para resolver o problema do reduzido número de VLAN IDs disponíveis, e da coordenação entre prestadores de serviços e os gestores de redes privadas, foi criada uma nova *tag*, também com 12 bits, a S-VLAN (Service Provider-VLAN), que é usada pelo prestador de serviços para identificar o serviço a que a *frame* do cliente pertence e encaminhá-la através sua rede, enquanto a VLAN-TAG, renomeada para C-VLAN (*Costumer-VLAN*) é usada dentro da rede privada do cliente [22]. Este empilhar de *tags*, faz com que esta solução seja também chamada de *Q-in-Q*, ou *stacked VLANs*.

DA	SA	C-VLAN	S-VLAN	Payload	FCS
----	----	--------	--------	---------	-----

Figura 2.9: *Frame* usada pelo protocolo 802.1AD, Provider Bridges

Os equipamentos encarregados de inserir e retirar as S-TAGs e que servem de fronteira entre as redes dos clientes e a rede do operador chamam-se Provider Edge Bridges. E os equipamentos que fazem uso da S-VLAN para encaminhar o tráfego chamam-se Provider Bridges.

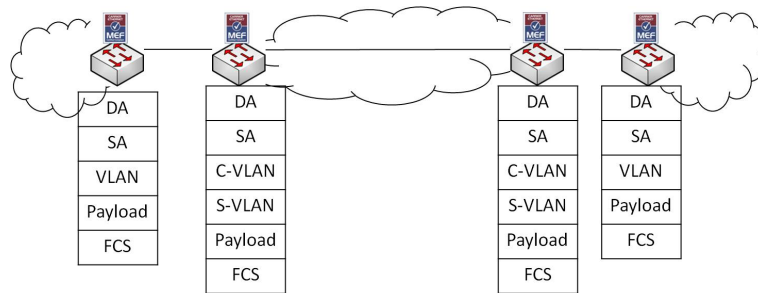


Figura 2.10: Aplicação da tecnologia Provider Bridges.

A hierarquia desta tecnologia está na imagem 2.10. O funcionamento desta tecnologia é o seguinte:

1. O equipamento e a topologia das redes privadas, permanecem inalterados.
2. A PEB detecta quando uma *frame* é endereçada para fora da rede privada, e quando assim acontece acrescenta a S-VLAN, e encaminha a *frame* para a rede do prestador de serviços onde é encaminhada com base na S-VLAN
3. Dentro da rede do prestador de serviços, a *frame* é encaminhada pelas PB de acordo com a sua S-VLAN
4. Analogamente quando uma *frame* chega a uma PEB vinda da rede do prestador de serviços e está endereçada para dentro da rede privada, a UNI retira a S-TAG para as bridges do cliente conseguirem encaminhar a *frame*.

Este empilhar de VLANs resolve o problema de coordenação de VLAN-IDs entre o prestador de serviços e os gestores das redes privadas mas não resolve nenhum dos outros problemas de aplicar a tecnologia Ethernet às redes de transporte.

2.6.2 IEEE 802.1AH - Provider Backbone Bridges

Para resolver os problemas de escalabilidade provocados por tabelas de endereços MAC demasiado grandes para serem pesquisadas em tempo útil, por um número insuficiente de VLANs que é possível criar com uma *tag* de 12 bits, e para distinguir os domínios das redes privadas e da rede de transporte foi necessário alterar novamente o cabeçalho Ethernet. A tecnologia que especifica estas alterações é a norma IEEE 802.1AH - Provider Backbone Bridges. Se a norma 802.1AD empilha VLAN-IDs para distinguir os VLAN-IDs dos clientes e do prestador de serviços, a norma 802.1AH empilha endereços MAC, a *frame* desta norma está na figura 2.11, e a rede, ou secção da rede que usa tecnologia é chamada de Provider Backbone Bridge Network (PBBN).

Frame

B-DA	B-SA	B-VID	I-SID	C-DA	C-SA	C-VLAN	Payload	FCS
------	------	-------	-------	------	------	--------	---------	-----

Figura 2.11: Provider Backbone Bridge *frame*.

Esta tecnologia acrescenta 4 *tags* à *frame* 802.1AD. Estas *tags* são:

- B-DA: Esta *tag* com 48 bits guarda o endereço de destino na PBBN.
- B-SA: Esta *tag* com 48 bits guarda o endereço da porta onde a *frame* entrou na PBBN
- I-SID: Esta *tag* com 24 bits, é única dentro da rede do operador, e identifica uma comunidade de interesse implementada como um domínio de broadcast[10]. É atribuído I-SID a cada instância de serviço do cliente [9]. Esta *tag* é visível apenas para as BEB, sendo transparente para as BCB.
- B-VID: Esta *tag* com 12 bits identifica a VLAN na rede do operador [10]. As B-VLANs são partições da PBBN que actuam como classes de serviço.

Nota: Os endereços de entrada e saída no cabeçalho PBB, são independentes dos endereços de destino e origem no cabeçalho Ethernet. Aliás, como o são dois domínios separados, endereços MAC atribuídos a equipamentos PBB podem sobrepor-se a endereços já atribuídos a equipamentos LAN [22].

Equipamentos

Para lidar com as *tags* introduzidas esta tecnologia introduz dois tipo de bridges novos, que são:

- Backbone Edge Bridge (BEB): Estas bridges estão na extremidade da PBBN, e são responsáveis por adicionar e retirar as *tags* PBB às *frames*. As BEB, têm dois componentes: o I-Component e o B-Component. Quando uma *frame* chega a uma BEB com destino à PBBN o I-Component mapeia o S-TAG em I-SID e o B-Component mapeia o C-TAG em B-VID e adiciona o cabeçalho PBB [6]. Quando o sentido da *frame* é o oposto, o funcionamento da BEB é invertido. O B-Component de uma BEB encaminha as *frames* com base nos endereços da PBBN, e o I-Component encaminha as *frames* com base nos endereços da PBN, assim estes componentes precisam de conhecer os endereços MAC quer das bridges da PBBN, quer das redes locais a que estejam ligadas, quer estas redes sejam PBN ou redes de clientes [6].
- Backbone Core Bridge (BCB): Estas bridges estão no *core* da PBBN, e fazem o encaminhamento das *frames* no *core* da PBBN, por isso apenas aprendem os endereços das restantes BCB e das BEB. Estas bridges têm um funcionamento análogo às PEB da tecnologia 802.1AD PB tanto que estrutura e a semântica da B-VID é igual à da S-TAG, e assim estas bridges não têm de estar cientes da PBBN, podendo até ser usadas bridges PB no *core* da PBBN [6]. Estas bridges o encaminham as *frames* usando o tuplo : <B-DA, B-VID>.

Conclusão

Com as *tags* inseridas na *frame* a tecnologia 802.1AH, consegue resolver os problemas de escalabilidade relativos ao número de serviços prestados, porque a I-SID permite, em teoria 2^{24} serviços, e 2^{12} classes de serviço. O problema de tabelas de endereçamento muito grandes, e a falta de distinção entre os domínios das redes dos clientes e do operador, foi resolvido com o empilhar de endereços MAC.

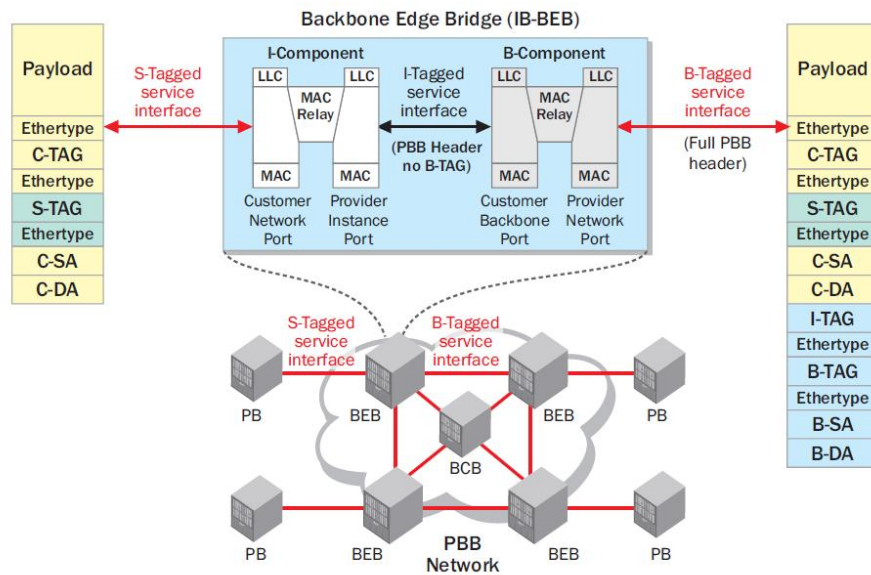


Figura 2.12: Esquema de uma Backbone Edge Bridge. Imagem disponível em [6]

2.6.3 IEEE 802.1Qay - Provider Backbone Bridges - Traffic Engineered Funcionamento

Usando as tecnologias descritas anteriormente, é possível aplicar engenharia de tráfego a uma zona da rede, esta engenharia de tráfego é feita com recurso a uma nova tecnologia: IEEE 802.1Qay - Provider Backbone Bridges. Esta tecnologia não introduz alterações no cabeçalho das *frames*, mas usa a premissa que "ignorando algumas funcionalidades Ethernet, o equipamento Ethernet é capaz de encaminhamento orientado à ligação"[22]. Com esta tecnologia o protocolo *Spanning Tree* é abandonado em favor de uma entidade de gestão externa [14] responsável por criar ligações ponto-a-ponto na rede PBB chamadas Ethernet Switched Paths ao preencher as *MAC filter tables* das bridges.

O primeiro passo para usar esta tecnologia é configurar um conjunto de B-VIDs retirados ao controlo do STP. Um B-VID que pertença a este conjunto é chamado de ESP-VID e é associado a uma valor especial do MSTP que indica que este VID específico não está sob o controlo de nenhum STP [11]. A alocação de *frames* é feita nas IB-BEBs, e o B-SA das *frames* é o endereço MAC das portas Ethernet do cliente que têm capacidade para mapear ESP's. Estes endereços MAC são designados por ESP-MACs, e estas portas são designadas por Customer Backbone Ports, ou CBPs[11].

A entidade de gestão externa, cuja relação com a rede está na figura 2.13, é responsável por:

- Auto-descobrir a topologia da rede [5]
- Criar árvores, ou ligações ponto-a-ponto através da rede, partindo de cada CBP pertencente à região PBB-TE ligando um conjunto de CBPs pertencentes a essa região [11].

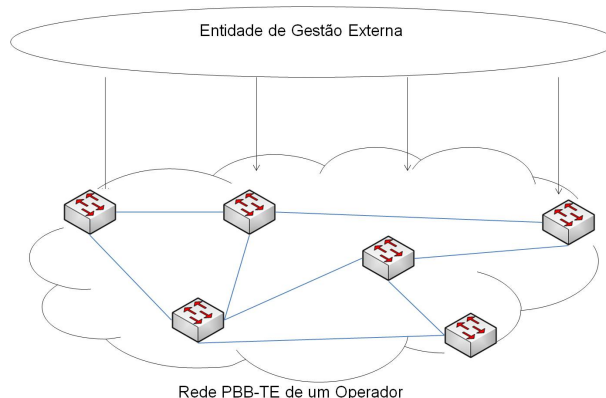


Figura 2.13: Relação entre o espaço PBB-TE e a entidade de gestão.

- Preencher as MAC *filter tables* das bridges pertencentes a essas árvores com entradas para os CBP DA e ESP-VIDs.
- Gerir a largura de banda de cada ESP ao longo de cada árvore [11].
- Para cada CBP que faz parte da região PBB-TE, provisionar uma ou mais árvores de que fornecem caminhos de retorno [11].

Neste contexto cada um dos ESPs é definido por um tuplo do tipo [11]:

<ESP-MAC SA, ESP-MAC DA, ESP-VID >

Em que :

- ESP-MAC SA: Endereço da CBP que está ligada à IB-BEB que é a raiz do ESP.
- ESP-MAC DA: Endereço que identifica a CBP de destino.
- ESP-VID: É um valor que distingue diferentes ESPs. Apenas podem ser atribuídos valores de ESP alocados à região PBB-TE.

Na norma PBB-TE estão definidos 2 tipos de ligação:

- Ligação ponto-a-ponto, constituída por dois ESPs, cada um fornecendo conectividade em cada sentido [11].
- Ligação ponto-a-multiponto, constituída por um ESP ponto-a-multiponto, e ESPs unidirecionais ponto-a-ponto mapeados ao longo as folhas da árvore[11].

Apesar de cada ESP ser identificado por um tuplo do tipo <ESP-MAC DA, ESP-MAC SA, ESP-VID> apenas o par <ESP-MAC DA, ESP-VID> é usado no encaminhamento das *frames*. Usar esta combinação permite reutilizar os B-VIDs atribuídos à região PBB-TE [22].

O comportamento de uma bridge que suporte PBB-TE é o seguinte:

- Não fazer *flooding* de *frames* cujo destino <ESP-MAC DA, ESP-VID> é desconhecido, quando isso acontece as *frames* são descartadas [11].
- Não adicionar entradas na tabela de encaminhamento através de *frames* recebidas. Esta tabela é preenchida e actualizada pela entidade externa de gestão [11].

- Não usar nenhum STP na região PBB-TE, toda a informação de encaminhamento é fornecida pela entidade externa de gestão [11].

Numa rede onde está implementada esta tecnologia em paralelo com VLAN, o estabelecimento de ESPs e o policiamento de tráfego à entrada dos ESPs é uma condição necessária, mas não suficiente para garantir que o tráfego na rede seja tratado com a QoS desejada, para garantir que essa QoS seja respeitada o tráfego PBB-TE deve ter sempre prioridade sobre o tráfego VLAN[11].

2.7 Solução proposta em conjunto por IETF/ITU-T

A solução proposta em conjunto por IETF/ITU-T, é um protocolo chamado MPLS-Transport Profile, e baseia-se no protocolo MPLS, este protocolo tradicionalmente opera entre as camada 2 e 3 do modelo OSI [16], e é um protocolo amplamente usado por operadores no encaminhamento de redes de pacotes. O protocolo MPLS é operado por dispositivos de camada 3, e transparente para os dispositivos de camada 2. O desafio para o protocolo MPLS-TP é modificar o protocolo MPLS para que este opere na camada de ligação de dados. O funcionamento do protocolo MPLS-TP é até certo ponto análogo ao funcionamento do protocolo MPLS, o mesmo tipo de pacote é usado, com a diferença exposta na figura 2.14, a diferença é que a RFC5462 [1], muda a denominação e a função do campo "EXP" do pacote MPLS tradicional para "Traffic Class (TC)" e este campo passa a transportar informação sobre a CoS do pacote, em [7] são analisadas várias formas de interpretar a informação contida neste campo.

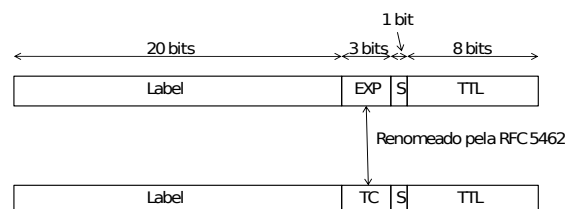


Figura 2.14: Cabeçalhos usados pelos protocolo MPLS e MPLS-TP

Apesar do esforço das entidades envolvidas, o protocolo MPLS-TP não está ainda totalmente definido, há ainda discussão sobre algumas RFCs que ainda não entraram em vigor. No entanto, todas as RFCs, definitivas e ainda em discussão estão disponíveis no site do IETF.

2.7.1 Diferenças entre MPLS e MPLS-TP

Ao nível de funcionamento existem 4 diferenças principais entre os protocolos MPLS e MPLS-TP:

- O protocolo MPLS-TP não implementa *Penultimate Hop Popping* [13], no protocolo MPLS tradicional era possível, para não sobrecarregar os LERs, desencapsular os pacotes no último LSR que o pacote passava. Esta funcionalidade conduz à perda de contexto, logo no protocolo MPLS-TP foi desactivada.

- O protocolo MPLS-TP não suporta *Load Balancing* [13], no protocolo MPLS as rotas entre os LSRs eram definidas manualmente, ou usando um protocolo de encaminhamento, podia acontecer que entre dois LSRs houvesse mais do que um caminho com o mesmo custo e se tal acontecesse os LSRs podiam distribuir o tráfego entre ambas as rotas, no protocolo MPLS-TP esta funcionalidade está desactivada. Pode acontecer que na rede haja *Load Balancing*, mas tal deve ser transparente para o protocolo MPLS-TP.
- No protocolo MPLS-TP um LSP bidireccional entre dois LSRs é obrigatoriamente congruente, isto quer dizer que o caminho percorrido pelos pacotes enviados por cada LSR deve ser o mesmo [19]. Como cada dispositivo que corra um IGP toma decisões de encaminhamento com base apenas na informação que possui, no caso do protocolo MPLS podia acontecer que os pacotes enviados em cada direcção percorressem caminhos diferentes, isso torna a monitorização de tráfego muito difícil, por isso no protocolo MPLS-TP foi tomada a decisão de forçar os LSPs bidireccionais a serem congruentes.
- No protocolo MPLS os pacotes eram encaminhados tendo em conta o seu destino, por isso pacotes pertencentes a vários LSPs encaminhados por um router para o mesmo destino, podiam ser encaminhados com a mesma *label*, tal facto torna impossível monitorizar a ligação entre um extremo e o outro, por isso essa funcionalidade foi desactivada no protocolo MPLS-TP

2.7.2 Aplicação de MPLS-TP a Carrier Ethernet

Se os problemas que aplicar a tecnologia Ethernet às redes de transporte enumerados em 2.5 forem recordados, pode ver-se que o protocolo MPLS-TP tem uma resposta para todos eles:

- Escalabilidade: O protocolo MPLS-TP usa routers e não switches para separar as redes, isto implica que cada dispositivo não tem de aprender os endereços MAC de todos os outros dispositivos na rede. No entanto pode ser por vezes necessário transportar a *frame* Ethernet através da rede MPLS-TP, neste caso a RFC4448[18] define a forma de encapsular e transportar *frames* Ethernet em pacotes MPLS. O problema do número de VLANs não se coloca porque o protocolo MPLS-TP não usa VLANs, identifica serviços com base em *labels* com 20 bits o que permite oferecer 1 milhão de serviços, e estas *labels* são passíveis de ser empilhadas. Se por exemplo um operador empilhar duas *labels* pode oferecer $2^{20} * 2^{20} = 2^{40}$ serviços. O problema das *frames* com destino desconhecido também não se coloca, porque à entrada da rede, nos LSRs os pacotes são filtrados, entre os que são encapsulados e entram na rede MPLS-TP, e os que são descartados. Os pacotes que entram na rede estão sujeitos, a não ser o caso de haver dispositivos mal configurados manualmente, a um processo de monitorização que garante que são entregues no LSR correcto.
- Protecção: A RFC5654 [19], define que para uma topologia em anel, o protocolo MPLS-TP deve fornecer um tempo de recuperação equivalente ao SDH, isto é a rede recupera de falhas em 50ms.
- QoS: A RFC RFC5654 [19], define como requisitos de QoS em redes MPLS-TP, o suporte a serviços diferenciados, o provisionamento e garantia de cumprimento de diferentes SLAs, o suporte de serviços sensíveis a jitter e atrasos, a garantia de acesso

justo aos recursos da rede, e o suporte e um esquema flexível de alocação de largura de banda. Estes requisitos permitem aos operadores oferecer aos clientes a garantia que os parâmetros acordados para os serviços oferecidos são cumpridos.

Capítulo 3

Modelo eficiente de distribuição de IPTV

O conceito de IPTV é muito vago, diversos organismos têm definições, ou participaram no esforço de definição do conceito de IPTV[20]:

- DSL Forum (actualmente conhecido como Broadband Forum)
- ETSI - European Telecommunications Standards Institute
- IPDR - Internet Protocol Detail Record Organization têm definições para IPTV
- ISMA - Internet Streaming Media Alliance
- DVB-IPI
- WirelessHD Consortium
- Moving Pictures Experts Group
- Digital Video Broadcasting Project

No entanto a definição de IPTV mais aceite é a definição proposta pelo ITU:

”IPTV é definido por distribuição de serviços multimédia tais como televisão/ video/ audio/ texto/ gráficos/ dados sobre redes IP geridas para prestar o nível necessário de qualidade de serviço e experiência, segurança, interactividade e confiança ”

Do ponto de vista dos clientes IPTV é um serviço que além do serviço tradicional de televisão, fornece serviços de Video-On-Demand (VoD) e televisão interactiva. Resumidamente as características de IPTV são [20]:

- Suporte televisão interactiva: A comunicação bidireccional permite aos operadores fornecer serviços interactivos
- ”Time shifting”: IPTV permite a gravação de TV e posterior visionamento

- Requisitos de baixa largura de banda: Em vez de enviar todos os canais para todos os utilizadores, os operadores podem enviar para cada utilizador o canal que ele pretende ver, ou criar sessões de multicast quando necessário
- Acessível em vários dispositivos: A visualização de IPTV não está limitada a televisões, os clientes podem usar os computadores, ou dispositivos móveis para aceder a conteúdos IPTV

Tecnicamente o conceito de IPTV não está definido, nem ao nível de aplicação, que neste caso são os *codecs* necessários para poupar largura de banda, e que todos os dispositivos na mesma rede que necessitam de interpretar os fluxos de dados de IPTV como sinais de televisão têm de usar os mesmos *codecs*

3.1 Modelo de distribuição de IPTV

Esta dissertação vai ter como objectivo modelar um sistema de IPTV usando como referência o modelo proposto em [3]. Este modelo atribui um custo ao envio de um canal em multicast, e um custo ao envio de um canal unicast. E assume que o custo de enviar um canal em multicast é maior que o custo de enviar um canal em unicast. Mas como o envio de um canal em unicast para cada cliente desperdiça muita largura de banda, a solução é um modelo misto que combine o envio de canais em unicast e multicast. Assim o modelo de um sistema de IPTV segue o esquema da figura 3.1. Este esquema hierárquico tem 3 camadas:

- Camada IPTV: Esta camada define quais os canais enviados em multicast e em unicast, segundo o modelo a apresentar a seguir.
- Camada Carrier Ethernet: Esta camada é responsável pelas ligações extremo-a-extremo
- Camada 1 (Física): Responsável por interagir com o hardware

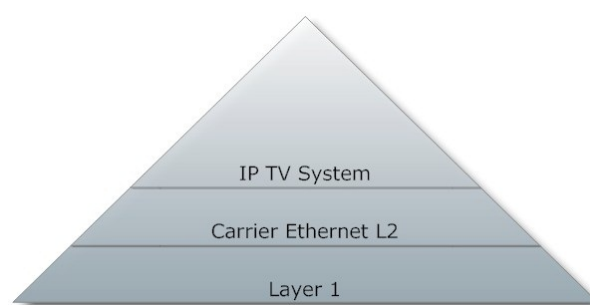


Figura 3.1: Camadas do sistema de IPTV

3.2 Modelar Sistema de IPTV

Segundo [3] por vezes enviar um canal em multicast consome mais recursos do que enviar em unicast. Assim o modelo proposto para otimizar o uso dos recursos da rede é o seguinte:

Vamos considerar que um operador quer oferecer um pacote com K canais. Estes podem ser ordenados pela popularidade (probabilidade de um utilizador estar a ver o canal). Há estudos [25] que mostram que a popularidade dos canais segue uma distribuição de Zipf, o vector Π guarda estas probabilidades, que são calculadas de acordo com a fórmula:

$$\pi_k = dk^{-\alpha}, \text{ para } k = 1, 2, \dots, K \quad (3.1)$$

O comportamento dos utilizadores é modelado da seguinte maneira: cada utilizador está activo (a ver um canal) com probabilidade a , e está inactivo com probabilidade $1-a$. Como a decisão de cada utilizador estar ou não activo é independente do comportamento dos outros utilizadores, este comportamento obedece às regras de uma variável de Bernoulli com parâmetro $p = a$. Como o comportamento de cada utilizador é uma variável de Bernoulli o comportamento do conjunto de utilizadores pode ser modelado como uma distribuição Binomial (Anexo C). O comportamento dos utilizadores tem influência no parâmetro d da distribuição de Zipf que modela a popularidade dos canais, desde que esteja activo, um utilizador tem obrigatoriamente de estar a ver um dos K canais, assim o somatório da probabilidade de cada canal estar a ser visto é obrigatoriamente 1. Assim, o parâmetro d pode ser calculado como:

$$\sum_{k=1}^K \pi_k = 1 \Leftrightarrow \sum_{k=1}^K dk^{-\alpha} = 1 \Leftrightarrow d \sum_{k=1}^K k^{-\alpha} = 1 \Leftrightarrow d = \frac{1}{\sum_{k=1}^K k^{-\alpha}}$$

3.3 Cenário dinâmico e cenário estático

Propomos dois cenários para um melhor aproveitamento dos recursos da rede, um cenário estático, e um cenário dinâmico.

- **Cenário Estático:** No cenário estático são escolhidos os M canais mais populares e esses são sempre enviados em multicast, e os restantes $(K - M)$ canais são enviados em unicast conforme são requisitados. O maior objectivo neste caso é calcular o número M óptimo de que permita otimizar os recursos da rede.
- **Cenário Dinâmico:** No cenário dinâmico o envio de um canal em multicast ou broadcast está dependente do número de utilizadores a ver o canal. Assim enquanto a largura de banda requerida para transmitir o canal para em unicast for inferior à necessária para transmitir o canal em multicast, o canal transmitido em unicast, quando o número de utilizadores faz aumentar essa largura de banda até ao limiar em que é mais eficiente o envio do canal em multicast, nesse momento canal passar a ser transmitido em multicast. Este facto implica que num dado momento e dependendo dos utilizadores activos, da popularidade dos canais num dado momento todos os canais podem estar a ser transmitidos em multicast, ou em unicast. Neste caso o maior objectivo é encontrar o limiar a partir do qual é mais eficiente enviar os canais em multicast.

3.4 Abordagem Exacta

Com uma abordagem exacta ao sistema de IPTV pretende-se saber com exactidão qual é o comportamento dos utilizadores, e os recursos da rede utilizados. Tendo em conta os

parâmetros da tabela 3.1, para obter uma abordagem exacta do modelo é preciso ainda definir o vector C como o vector com entradas c_k ($k = 1, 2, \dots, K$) que correspondem ao número de utilizadores a ver o canal k . Deste vector é fácil obter o número de utilizadores inactivos como N menos a soma de todas as entradas c_k . É também definida a variável $n_a = c_1 + c_2, \dots, + c_K$, que é o número de utilizadores activos num dado momento.

3.4.1 Cenário Estático

Neste cenário o número de canais transmitidos em multicast está fixado em M , mas n_u i.e. o número de canais transmitidos em unicast é uma variável aleatória. Dependendo da forma como os utilizadores estão ou não activos, e a sua distribuição pelos canais, o número de canais enviados pelo sistema tanto podem ser apenas os M canais enviados em multicast, como num caso extremo são enviados M canais em multicast, que não estão a ser vistos por nenhum utilizador, e, simultaneamente são enviados N canais unicast (um por utilizador). Assim é necessário saber a probabilidade $Pr[n_u = i]$ i.e. a probabilidade do número de canais unicast ser igual a i . Esta probabilidade depende:

- Do número de utilizadores activos
- Do vector Π que influencia a distribuição de popularidade dos canais
- Do valor de M que influencia o número de canais que podem ser enviados em unicast

A probabilidade que se quer calcular é: tendo n_a utilizadores activos, qual a probabilidade de serem transmitidos n_u canais em unicast. Para calcular esta probabilidade para todas as possibilidades de utilizadores activos (o número de utilizadores activos varia entre 0 e N), soma-se a probabilidade de haver n utilizadores activos e serem enviados i canais em unicast:

$$Pr[n_u = i] = \sum_{n=0}^N Pr[n_u = i \wedge n_a = n] \quad (3.2)$$

Com os dados que dispomos é impossível calcular $Pr[n_a = n \wedge n_u = i]$ directamente, assim, usamos a expressão da probabilidade condicional para chegar a uma expressão que conseguimos calcular. Os passos para conseguir calcular a probabilidade de haver n utilizadores activos e serem enviados i canais em unicast, partindo da expressão para a probabilidade condicional são:

$$\begin{aligned} Pr[n_u = i | n_a = n] &= \frac{Pr[n_u = i \wedge n_a = n]}{Pr[n_a = n]} \\ Pr[n_u = i \wedge n_a = n] &= Pr[n_u = i | n_a = n] \cdot Pr[n_a = n] \end{aligned} \quad (3.3)$$

Como é possível calcular $Pr[n_u = i | n_a = n]$ e $Pr[n_a = n]$, já podemos calcular $Pr[n_u = i]$. De seguida calculamos os valores das expressões anteriores.

$$\begin{aligned} Pr[n_a = n | n_u = i] &= \frac{n!}{(n-i)!i!} (1 - P_u)^{n-i} (P_u)^i \\ Pr[n_a = n] &= \frac{N!}{(N-n)!n!} (1 - a)^{N-n} a^n \end{aligned} \quad (3.4)$$

Assim:

$$\begin{aligned}
Pr[n_u = i] &= Pr[n_a = n | n_u = i] \cdot Pr[n_a = n] \\
&= \frac{n!}{(n-i)!i!} (1 - P_u)^{n-i} (P_u)^i \cdot \frac{N!}{(N-n)!n!} (1-a)^{N-n} a^n \\
&= \frac{N!}{(N-n)!} (1-a)^{N-n} a^n \frac{(P_m)^{n-i} (P_u)^i}{(n-i)!i!}
\end{aligned} \tag{3.5}$$

Nas expressões P_u e P_m são respectivamente a probabilidade de um utilizador escolher ver um canal unicast ou um canal multicast. Como M é um valor fixo, a probabilidade P_m é dada pela soma das probabilidades de um utilizador ver algum dos M canais enviados em multicast. E P_u é a probabilidade complementar de P_m . Estas expressões são:

$$\begin{aligned}
P_m &= \sum_{k=1}^M \pi_k \\
P_u &= 1 - P_m
\end{aligned} \tag{3.6}$$

Na figura 3.2 está a representação da função de densidade de probabilidade de $Pr[n_u = i]$ para um cenário com 25 canais ($K=25$), em que 15 são sempre enviados em multicast ($M=15$), com 170 utilizadores ($N=17$), sendo que a probabilidade de cada utilizador estar activo é 0.7 ($a=0.7$), e com parâmetros $\alpha=0.5$, e $\beta=2$.

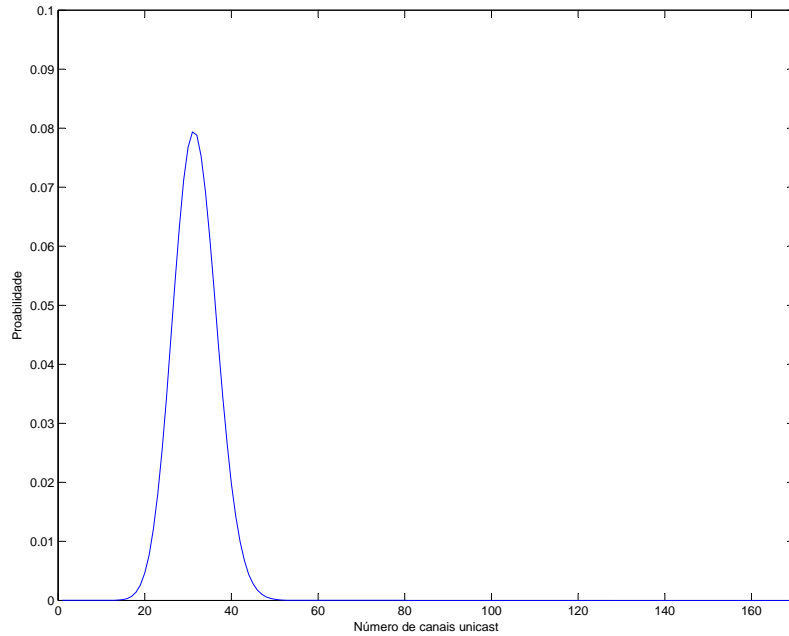


Figura 3.2: Função de distribuição de probabilidade de $Pr[n_u = i]$, isto é, a probabilidade de haver i canais enviados em unicast

3.4.2 Cenário Dinâmico

Para conseguir uma aproximação exacta no cenário dinâmico é necessário definir um conjunto de vectores B , contendo entradas b_k ($k = 1, 2, \dots, K$). Cada um destes vectores representa uma maneira possível, e diferente de n_a utilizadores activos se distribuírem pelos K canais, e o conjunto de todos os vectores B representa todas as maneiras possíveis de todos n_a utilizadores activos (de nenhum utilizador activo a N utilizadores activos) se distribuírem pelos K canais. Segundo [24] o número de vectores B possíveis é dado por 3.7

$$C(K + N, N) = \frac{(N + K)!}{N!K!} \quad (3.7)$$

Para calcular com exactidão as probabilidades de um determinado canal ser transmitido em multicast ou unicast, há que calcular a probabilidade ω_B associada a cada vector B , esta probabilidade depende de:

- Da probabilidade de cada utilizador estar activo
- Da distribuição da popularidade dos canais
- Do número de utilizadores
- Da distribuição de utilizadores representada pelo vector B .

O vector ω_B representa a probabilidade da distribuição presente no vector B ser a distribuição real dos utilizadores activos pelos K canais. A expressão para calcular ω_B é dada por:

$$\omega_b = Pr[c_1 = b_1, \dots, c_K = b_K]$$

Novamente, não é possível calcular directamente esta grandeza, e é necessário recorrer à expressão da probabilidade condicional. Assim, há que trabalhar a expressão até atingirmos expressões que consigamos calcular.

$$\begin{aligned} Pr[c_1 = b_1, \dots, c_K = b_K | n_a = n] &= \frac{Pr[c_1 = b_1, \dots, c_K = b_K]}{Pr[n_a = n]} \\ Pr[c_1 = b_1, \dots, c_K = b_K] &= Pr[c_1 = b_1, \dots, c_K = b_K | n_a = n] \cdot Pr[n_a = n] \end{aligned} \quad (3.8)$$

Desta maneira temos duas expressões que podemos calcular. $Pr[n_a = n]$ segue uma distribuição binomial (uma explicação sobre a distribuição binomial está no Anexo C), assim:

$$Pr[n_a = n] = \frac{N!}{(N - n)!n!} (1 - a)^{N - n} a^n$$

$Pr = [c_1 = b_1, \dots, c_K = b_K | n_a = n]$ segue uma distribuição multinomial (uma explicação sobre a distribuição multinomial está no Anexo C):

$$Pr = [c_1 = b_1, \dots, c_K = b_K | n_a = n] = \frac{n!}{b_1! \dots b_K!} \pi_1^{b_1} \dots \pi_K^{b_K}$$

Juntando as duas partes, para encontrar a expressão de ω_B :

$$\begin{aligned}
\omega_B &= \frac{N!}{(N-n)!n!} (1-a)^{N-n} a^n \frac{n!}{b_1! \dots b_k!} \pi_1^{b_1} \dots \pi_k^{b_k} \\
&= \frac{N!}{(N-n)!n!} (1-a)^{N-n} a^n n! \prod_{k=1}^K \frac{\pi_k^{b_k}}{b_k!} \\
&= \frac{N!}{(N-n)!} (1-a)^{N-n} a^n \prod_{k=1}^K \frac{\pi_k^{b_k}}{b_k!}
\end{aligned} \tag{3.9}$$

Neste cenário quer $Pr[n_m = i]$ quer $Pr[n_u = j]$, respectivamente a probabilidade de i canais serem enviados em multicast e a probabilidade de j canais serem enviados em unicast, são variáveis aleatórias calculadas segundo 3.10 e 3.11.

$$Pr[n_m = i] = \sum_B 1_{\{B: \text{multicast channels}=i\}} \omega_B \tag{3.10}$$

$$Pr[n_u = j] = \sum_B 1_{\{B: \text{unicast channels}=j\}} \omega_B \tag{3.11}$$

A função $1_{\{expr\}}$, definida por 3.12, é uma função que tem apenas dois resultados possíveis dependendo de "expr", caso "expr" seja uma expressão verdadeira então a função retorna o valor "1", caso "expr" seja uma expressão falsa a função retorna o valor "0". Assim, neste caso a função $\sum_B 1_{\{B: \text{multicast channels}=i\}}$ é a soma da probabilidade de cada vector B em que "i" canais são enviados em multicast. O número de canais multicast são calculados percorrendo cada vector B e contando o número de canais que estão a ser vistos por mais de β utilizadores.

$$1_{\{X\}} = \begin{cases} 1 & , \text{se } X \text{ verdadeiro} \\ 0 & , \text{se } X \text{ falso} \end{cases} \tag{3.12}$$

Na figura 3.3 está a representação das funções de densidade de probabilidade de $Pr[n_u = i]$ e $Pr[n_m = i]$ para um cenário com 6 canais ($K=6$), com 3 utilizadores ($N=3$), sendo que a probabilidade de cada utilizador estar activo é 0.7 ($a=0.7$), e com parâmetros $\alpha=0.5$, e $\beta=2$.

Esta simulação teve valores de N e K tão pequenos, porque o algoritmo exacto tem a desvantagem de obrigar a calcular, e percorrer todos os vectores B. O problema é que o número de vectores B aumenta de acordo com 3.7, o que faz com que rapidamente o número de vectores B seja tão grande que esta solução se torna impraticável. Na figura 3.4 está a evolução do número de vectores B para 10, 20, 30, 40, 50 e 100 utilizadores em cenários com 10, 15, 20, 25, 50 e 75 canais, nos gráficos pode ver-se que rapidamente o número de vectores B a calcular rapidamente dispara para valores que tornam esta abordagem muito pesada computacionalmente.

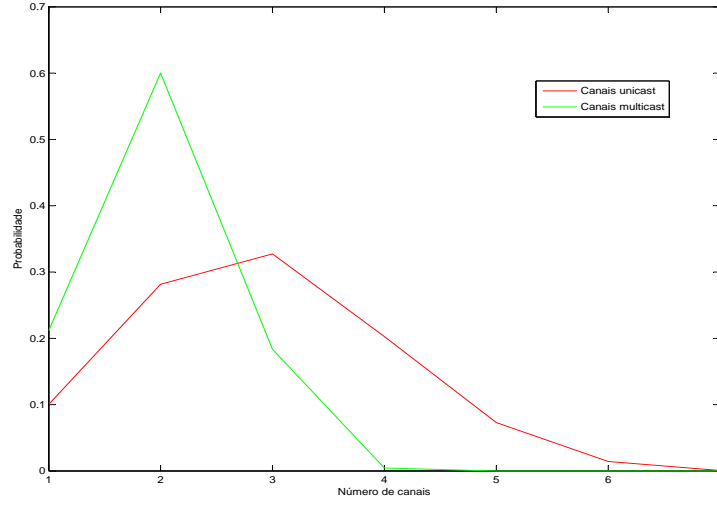


Figura 3.3: Funções de distribuição de probabilidade de $Pr[n_u = j]$ e $Pr[n_m = i]$, ou seja a probabilidade de haverem j canais unicast e i canais multicast

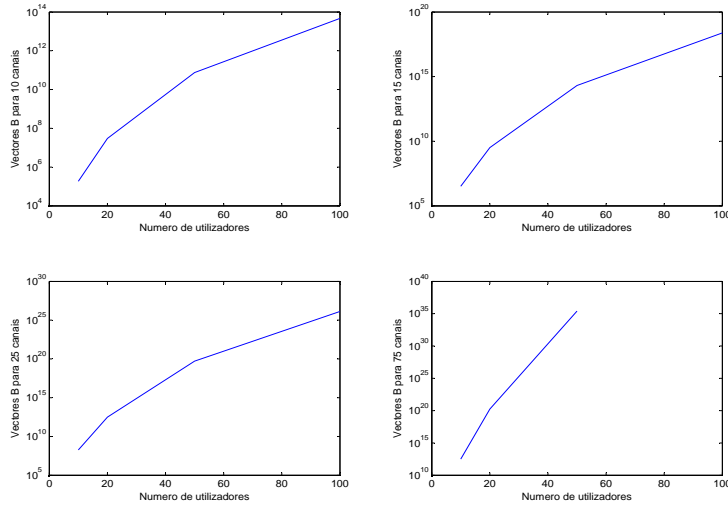


Figura 3.4: Variação da quantidade de vetores B. Nota: Nos gráficos é usada uma escala logarítmica.

3.5 Aproximação usando a distribuição Normal

Para contornar a desvantagem de a abordagem exacta ser muito pesada computacionalmente, pode assumir-se que as variáveis n_u e n_m podem ser aproximadas por uma distribuição normal. Esta forma de encarar o problema produz apenas aproximações, e não valores exactos,

mas é mais leve computacionalmente.

A função de densidade da distribuição normal é calculada usando a expressão 3.13 [26]. Para podermos calcular esta função, temos de calcular quer a média quer o desvio padrão da nossa distribuição. No nosso caso o que nos interessa é a média e o desvio padrão dos canais em unicast e multicast que são enviados.

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} \quad (3.13)$$

Neste caso, a média e o desvio padrão do número de canais a enviar em são definidas como:

- n_u : Número médio de canais unicast.
- n_m : Número médio de canais multicast.
- σ_u : Desvio Padrão dos canais unicast
- σ_m : Desvio Padrão dos canais multicast

Na figura 3.5 está a comparação entre as duas funções de densidade de probabilidade para o cenário estático apresentadas atrás. Pode constatar-se que a aproximação com a distribuição normal consegue, com elevado grau de precisão replicar o comportamento das funções de densidade de probabilidade do modelo.

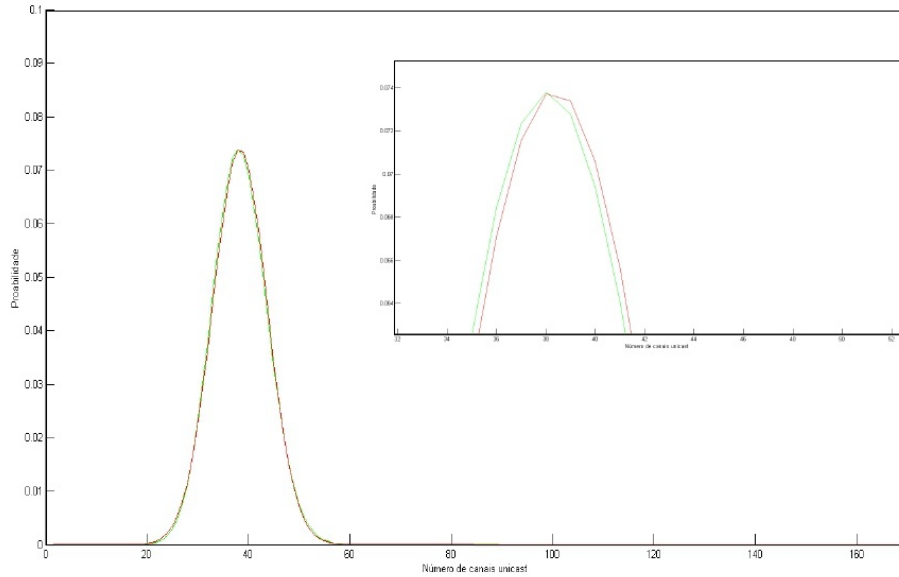


Figura 3.5: Comparação entre a abordagem exacta e a aproximação usando a distribuição normal.

3.5.1 Cenário Estático

No cenário estático, como $n_m=M$, apenas é necessário calcular a média e a variância de n_u . A média de n_u , designada por $E[n_u]$ calcula-se através da média de canais enviados em unicast sabendo o número de utilizadores activos, $E[n_u|n_a]$, e calculando depois a média de utilizadores activos, como mostrado em 3.14.

$$E[n_u] = E_{n_a}[E[n_u|n_a]] = E_{n_a}[n_a P_u] = N a P_u \quad (3.14)$$

$$Var[n_u] = E[n_u^2] - E[n_u]^2 \quad (3.15)$$

Como nós já conhecemos o valor de $E[n_u]$ de 3.14, resta calcular o valor de $E[n_u^2]$. O valor de $E[n_u^2]$, que é o momento de segunda ordem da variável, é calculado em 3.16.

$$\begin{aligned} E[n_u^2] &= E_{n_a}[E[(n_u)^2|n_a]] \\ &= E_{n_a}[n_a P_u(1 - P_u) + n_a^2 P_u^2] \\ &= n a P_u(P_m + (1 - a)P_u) + (N a P_u)^2 \end{aligned} \quad (3.16)$$

Assim, usando 3.15, e sabendo que o desvio padrão de uma variável aleatória é a raiz quadrada da variância, o desvio padrão de $E[n_u]$ é dado por 3.17:

$$\begin{aligned} \sigma &= \sqrt{E[X^2] - E[X]^2} \\ &= \sqrt{(n a P_u(P_m + (1 - a)P_u) + (N a P_u)^2) - (N a P_u)^2} \\ &= \sqrt{n a P_u(P_m + (1 - a)P_u)} \end{aligned} \quad (3.17)$$

A expressão para a largura de banda utilizada tem uma parte constante que depende de M e de β que é largura de banda ocupada pelos canais perpetuamente enviados em *multicast*, uma parte que varia com a probabilidades de os utilizadores estarem activos e de estarem a ver um canal que é enviado em *unicast*. A expressão $erfc^{-1}(P_{Block} \sqrt{N a P_u(1 - a P_u)})$ serve para corrigir o erro introduzido pela aproximação usando a distribuição normal.

Na figura 3.6 está a representação da função de densidade de probabilidade de $Pr[n_u = i]$ para um cenário com 25 canais ($K=25$), em que 17 são sempre enviados em multicast ($M=17$), com 170 utilizadores ($N=170$), sendo que a probabilidade de cada utilizador estar activo é 0.7 ($a=0.7$), e com parâmetros $\alpha=0.5$, e $\beta=2$.

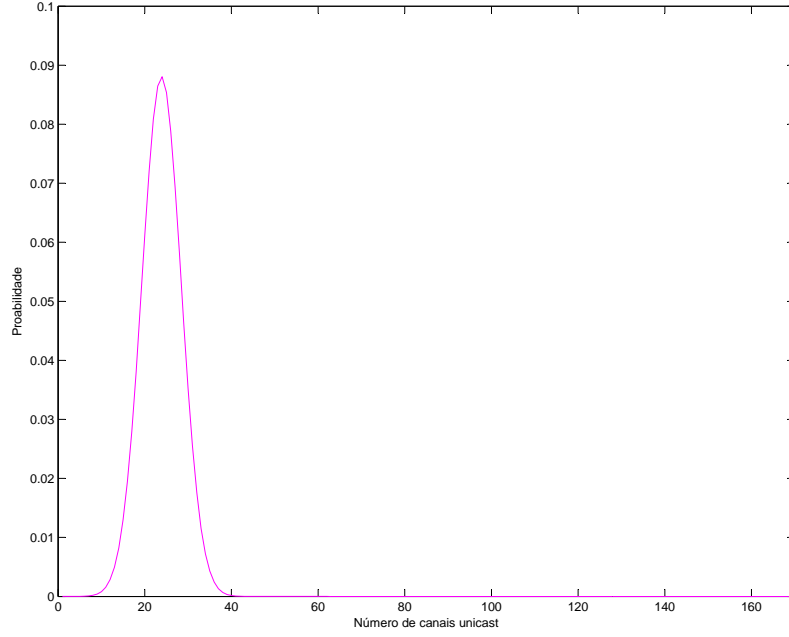


Figura 3.6: Função de distribuição de probabilidade de $Pr[n_u = i]$, isto é, a probabilidade de haver i canais enviados em unicast

3.5.2 Cenário Dinâmico

No cenário dinâmico quer o número de canais enviados em unicast, n_u quer o número de canais enviados em multicast n_m são variáveis aleatórias. Logo antes de achar o valor de r_D , há que achar o valor de $E[n_u]$ e de $E[n_m]$.

Neste cenário, o número de canais unicast, é dado pela soma de todos os utilizadores a ver canais unicast. Um canal unicast é um canal que tem um número de utilizadores sintonizados no intervalo $[1..\beta]$, este intervalo é definido pelo facto de um canal sem utilizadores sintonizados não ser transmitido, e, porque neste caso estamos a lidar com o cenário dinâmico, um canal com mais de β utilizadores sintonizados ser transmitido em multicast. A definição anterior dá-nos o número de canais enviados em unicast, mas, no caso de envio de canais em unicast, cada canal é enviado de forma independente para cada cliente, logo temos de saber quantos clientes estão sintonizados em canais unicast. Desta necessidade vem 3.19, nesta fórmula o número de canais unicast (n_u) é definido como a soma de todos os clientes que estão sintonizados num canal enviado em unicast, daí o vector C (que guarda o número de utilizadores sintonizados em cada canal) ser percorrido, e para cada canal, caso o número de utilizadores sintonizados esteja no intervalo $[1..\beta]$ soma-se esse número de utilizadores sintonizados ao número de canais enviados.

$$n_u = \sum_{k=1}^K \sum_{1 \leq i \leq \beta} i 1_{\{c_k=i\}} \quad (3.18)$$

Como a probabilidade da função indicadora, é a probabilidade do evento, podemos calcular $E[n_u]$ como:

$$E[n_u] = \sum_{k=1}^K \sum_{1 \leq i \leq \beta}^i Pr[c_k = i]$$

Ou seja $E[n_u]$, é a soma das probabilidades de cada canal ter entre 1 e β utilizadores. Esta probabilidade é calculada como $Pr[c_k = i]$, que é a probabilidade de o canal ter i utilizadores sintonizados, variando i entre no intervalo $[1..\beta]$ obtemos a probabilidade de o canal ser enviado em unicast. De seguida temos de calcular a probabilidade $Pr[c_k = i]$. Esta probabilidade também segue uma distribuição binomial, neste caso a experiência é descrita como a tomada de decisão de cada utilizador de estar activo (a ver um canal) e escolher um determinado canal. Assim, neste caso a probabilidade de sucesso do acontecimento, ou seja de o utilizador estar activo e a ver o canal k é dada por $a\pi_k$, e este exercício é repetido para cada um dos N utilizadores. Assim, em 3.19 está a expressão que define $Pr[c_k = i]$.

$$Pr[c_k = i] = \frac{N!}{i!(N-i)!} (1 - a\pi_k)^{N-i} (a\pi_k)^i \quad (3.19)$$

De seguida temos de calcular o desvio padrão de n_u , novamente calculamos o desvio padrão como a raiz quadrada da variância, e a variância é calculada segundo 3.23.

$$V[n_u] = \sum_{\substack{k=1 \\ k \neq l}}^K \sum_{l=1}^K \sum_{1 \leq i \leq \beta}^i \sum_{1 \leq j \leq \beta}^j (Pr[c_k = i, c_l = j] - Pr[c_k = i]Pr[c_l = j]) + \quad (3.20)$$

$$= \sum_{k=1}^K \left(\sum_{1 \leq i \leq \beta}^i Pr[c_k = i]^2 - \sum_{1 \leq i \leq \beta}^i \sum_{1 \leq j \leq \beta}^j Pr[c_k = i]Pr[c_k = j] \right) \quad (3.21)$$

Como estamos a lidar com o cenário dinâmico, em que os canais são enviados em multicast ou unicast conforme o número de utilizadores que os estão a ver, temos de prever também seguida vamos achar a média e o desvio padrão dos canais enviados em multicast. O número de canais enviados em multicast é dado pelo somatório de todos os canais que têm mais que β utilizadores sintonizados:

$$n_m = \sum_{k=1}^K 1_{\{c_k > \beta\}}$$

Usando a mesma propriedade que anteriormente, a probabilidade da função indicadora, é a probabilidade do evento, assim, podemos definir a $E[n_m]$ como a soma das probabilidades de haver mais de β utilizadores sintonizados em cada canal, ou como o acontecimento contrário a haver entre 0 e β utilizadores a ver um determinado canal.

$$\begin{aligned} E[n_m] &= \sum_{k=1}^K \sum_{i > \beta} Pr[c_k = i] \\ &= \sum_{k=1}^K 1 - \sum_{0 \leq i \leq \beta} Pr[c_k = i] \end{aligned} \quad (3.22)$$

E a variância de $E[n_m]$ é dada por:

$$\sum_{\substack{k=1 \\ k \neq l}}^K \sum_{l=1}^K \sum_{i > \beta}^i \sum_{j > \beta}^j (Pr[c_k = i, c_l = j] - Pr[c_k = i]Pr[c_l = j]) +$$

$$\sum_{k=1}^K \left(\sum_{i > \beta} Pr[c_k = i] \left(1 - \sum_{j > \beta} Pr[c_k = j] \right) \right)$$

Como r_D é dado por $r_D = n_u + \beta M$ A variância de r_D pode ser obtida como:

$$V[r_D] = E[r_D^2]E[r_D]^2 = V[n_u] + \beta^2 V[n_m] + 2\beta A \quad (3.23)$$

Na figura 3.7 está a representação da função de densidade de probabilidade de $Pr[n_u = j]$ e $Pr[n_m = i]$ para um cenário com 25 canais ($K=25$), em que 15 são sempre enviados em multicast ($M=15$), com 170 utilizadores ($N=17$), sendo que a probabilidade de cada utilizador estar activo é 0.7 ($a=0.7$), e com parâmetros $\alpha=0.5$, e $\beta=2$.

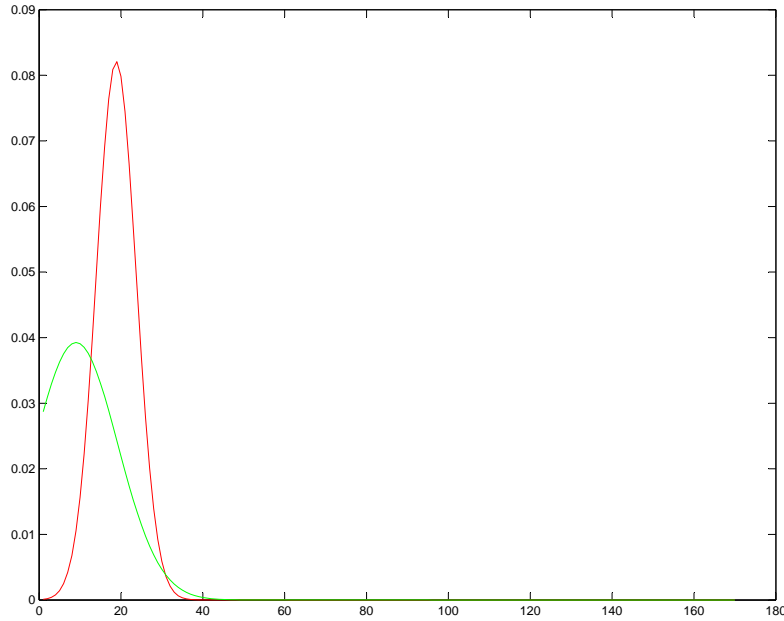


Figura 3.7: Funções de distribuição de probabilidade de $Pr[n_u = j]$ e $Pr[n_m = i]$, ou seja a probabilidade de haverem j canais unicast e i canais multicast

3.6 Optimização dos parâmetros do Modelo

Para um uso eficiente do modelo são vários os parâmetros que podem ser optimizados:

- No cenário estático, o valor de M , pode ser otimizado porque um valor de M grande demais significa que estão a ser enviados um canais multicast sem que esteja efectivamente a ser vistos, e um valor de M pequeno demais pode significar que o canal está a ser enviado em unicast para vários cliente, quando uma sessão multicast seria mais eficiente.
- No cenário dinâmico, o valor de β , pode ser otimizado para que o modo de transmissão dos canais alterne entre unicast e multicast de forma a utilizar os recursos da rede da forma mais eficiente.

3.6.1 Parâmetros do modelo

Os parâmetros do modelo estão na tabela 3.1.

Tabela 3.1: Parâmetros do modelo [3]

Parâmetro	Descrição
K	Número de canais
N	Número de utilizadores
a	Probabilidade de um utilizador estar activo
β	Relação entre a largura de banda requerida para o multicast de um canal e a largura de banda requerida para o unicast desse mesmo canal
Π	Vector com a popularidade dos canais
α	Parâmetro que controla a distribuição da popularidade dos canais

Capítulo 4

Resultados

Depois de ter um modelo estabelecido, há que o testar. Para testar o modelo, e obter resultados, foram criadas várias simulações em Matlab. Como a aproximação exacta descrita no capítulo anterior, é muito pesada, foi usada aproximação normal. A aproximação normal usada assume que:

$$R = R_M + R_U \quad (4.1)$$

Em que R é a carga que o sistema de IPTV impõe na rede, esta carga é medida em "unicast units", em que 1 "unicast unit" são os recursos necessários para transmitir um canal em unicast. Voltando a 4.1, R é igual à soma dos recursos necessários para enviar os canais em multicast, na expressão representados por R_M , mais os recursos necessários para enviar os canais em unicast, na expressão representados por R_U . A forma como estes recursos são calculados varia com os cenários.

4.1 Cenário Estático

No cenário estático o valor mais importante que importa otimizar é o valor de M , o número óptimo de canais a enviar em multicast. O valor de M pode ser otimizado de duas formas: pode ser calculado analiticamente ou, para cada cenário, pode ser procurado iterativamente de entre todas as possibilidades. O valor óptimo de M tem de ser calculado para cada cenário, porque o valor óptimo de M depende de vários parâmetros do sistema, desde logo o mais óbvio é o valor de K , porque, como já foi visto o valor de M varia entre 0 e K . Mas o valor de M depende, também, da distribuição da popularidade dos canais, neste caso em que usamos uma distribuição de Zipf, o valor óptimo de M depende intimamente do parâmetro α que controla a distribuição da popularidade dos canais, M depende ainda dos parâmetros a e N que controlam o número de utilizadores activos do sistema, o valor óptimo de M depende ainda do valor de β , porque este parâmetro controla o custo de enviar os canais em multicast em relação a enviá-los em unicast.

4.1.1 Calcular o valor de M

Em oposição a calcular o valor óptimo de M usando métodos analíticos, pode calcular-se, para cada cenário, a carga que diferentes valores de M impõem na rede, e escolher o valor de M que minimiza essa carga, além de calcular a carga imposta na rede por cada cenário, pode

também verificar-se como os diferentes parâmetros do sistema têm influência no valor óptimo de M . Neste cenário, as simulações assumem que:

$$R = \beta M + NaP_u \quad (4.2)$$

Nas simulações o valor de M é uma variável cujo valor pretendemos otimizar, e NaP_u já foi calculado no capítulo anterior.

4.1.2 Influência do valor de α

Vamos começar por definir 4 cenários em que varia o valor de α , a forma como o valor de α influencia a distribuição da popularidade dos canais está no anexo 3. Os parâmetros para os cenários estão na tabela 4.1.

	Cenário 1	Cenário 2	Cenário 3	Cenário 4
N	10000	10000	10000	10000
K	150	150	150	150
a	0.7	0.7	0.7	0.7
α	1.4	2	2.5	3
β	2	2	2	2

Tabela 4.1: Parâmetros dos cenários para iterativamente calcular o valor de M

Na figura 4.1 está a evolução da carga no sistema, para cada valor diferente de α , em função do valor de M que varia entre 0 e K . Para ajudar na leitura do gráfico está também presente a linha horizontal βM que representa o cenário em que $K = \beta M$ e a carga sobre a rede é constante independentemente do número de utilizadores activos, ou da distribuição da popularidade dos canais.

Pelos gráficos, nota-se a tendência para que, conforme a popularidade dos canais se concentra num número mais reduzido de canais, com o aumento de α , o número M óptimo de canais a enviar em multicast é mais reduzido. Este facto acontece porque se, para o mesmo número de clientes, estes se concentram mais num reduzido número de canais então menos clientes se distribuem pelos canais menos populares e o custo de enviar o canal em multicast compensa para um número menor de canais. Os resultados para estas distribuições de popularidade dos canais estão na tabela 4.2.

Pelos resultados, pode ver-se que para este cenário, o valor de α de 1,4, e também para valores menores que 1,4, a distribuição dos clientes pelos canais é tão uniforme que o valor óptimo de M é $M = K$, isto porque tendo $\beta = 2$ basta haver mais que dois clientes a ver um canal, para que enviar o canal unicast não seja mais eficaz que enviar o canal em multicast, e neste cenário com 10000 clientes, e $\alpha \leq 1.4$ há sempre mais de 2 clientes a ver qualquer canal.

4.1.3 Influência do número de utilizadores activos

Neste caso, vamos variar o valor do número de utilizadores activos, esta variação pode ser feita alterando dois parâmetros: o número de utilizadores no sistema, parâmetro N , ou

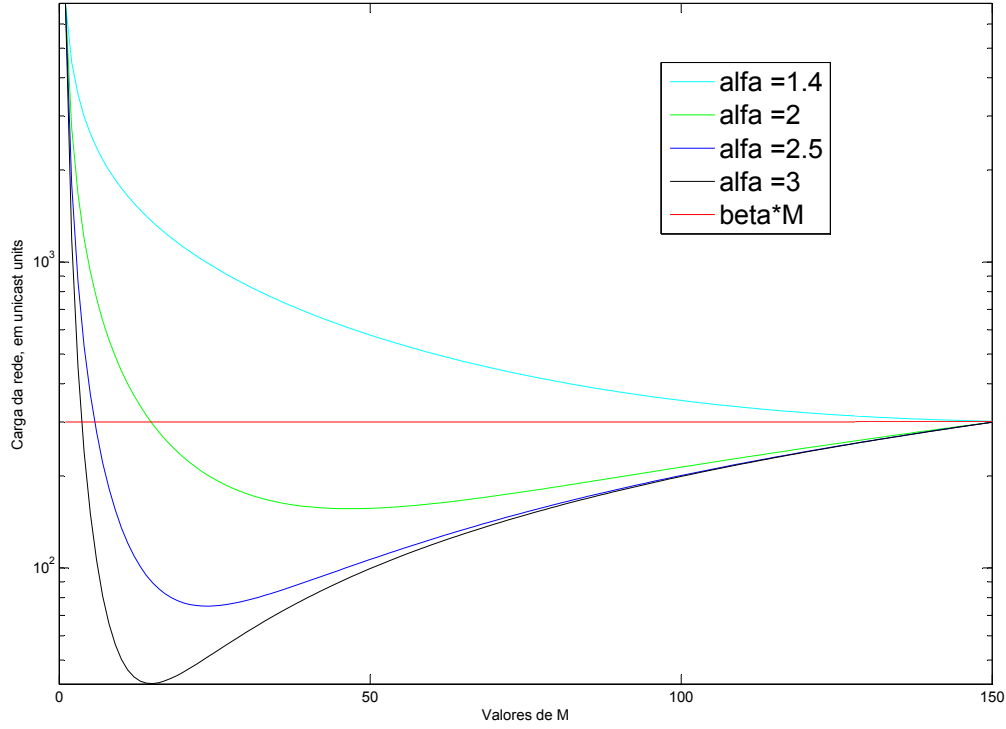


Figura 4.1: Evolução da carga no sistema, para diferentes distribuições de popularidade de canais

a probabilidade de estes estarem activos, parâmetro a . Neste caso vamos alterar apenas o parâmetro N . Os valores da simulação são os que estão na tabela 4.3.

Na figura 4.2, está a evolução da carga sobre o sistema dependendo do número de utilizadores activos. Pode ver-se que o aumento de utilizadores activos faz com que o valor óptimo de M também aumente. Este comportamento deve-se ao facto de que com o aumento de utilizadores activos, alguns dos canais menos populares são vistos por utilizadores suficientes para que o custo de os transmitir em multicast compense o custo de os transmitir em unicast para vários utilizadores.

Na tabela 4.4 estão os resultados para os valores óptimos de M para os cenários em que os valores de utilizadores activos variam. É possível ver que a variação do número de utilizadores influencia o número M óptimo de canais a enviar em multicast como a distribuição da popularidade dos canais.

4.2 Cenário Dinâmico

No cenário dinâmico o parâmetro que tem maior influência sobre a carga a que o sistema pode estar sujeito, é o parâmetro β , visto que todos os canais são enviados em unicast ou multicast dependendo apenas do valor de β . Neste cenário não faz sentido falar em optimização de parâmetros, visto que β depende da tecnologia usada, N , a e α dependem do

α	Carga Mínima	M óptimo
1.4	300	150
2	154	47
2.5	74	24
3	41	15

Tabela 4.2: Valores de M óptimos para diferentes distribuições de popularidade

	Cenário 1	Cenário 2	Cenário 3	Cenário 4
N	5000	15000	25000	35000
K	150	150	150	150
a	0.7	0.7	0.7	0.7
α	2.5	2.5	2.5	2.5
β	2	2	2	2

Tabela 4.3: Parâmetros dos cenários para iterativamente calcular o calor de M

cenário porque modelam o número de utilizadores activos, e a distribuição da popularidade dos canais, o valor de K também não pode ser optimizado porque por norma os operadores querem oferecer o maior número de canais possível aos clientes. Assim, resta perceber de que forma a variação destes parâmetros afecta o sistema. Para ter essa percepção testámos vários cenários, em que varia apenas um parâmetro, para assim ver a influência desse parâmetro. Nas simulações é assumido que:

$$R = \pi N a P_u \quad (4.3)$$

E partindo desta assumption são depois calculados os canais enviados em unicast e em multicast

4.2.1 Variação de β

O parâmetro β é a relação do custo entre enviar um canal em unicast, e o custo de enviar um canal em multicast. O valor de β depende da tecnologia que suporta o multicast. Para testar a influência de β na esforço exercido sobre a rede, correu-se uma simulação com os parâmetros da tabela 4.5. Nos teste o valor de β variou entre 1 que é caso em que o custo de enviar um canal em unicast é igual ao custo de enviar o canal em multicast, e o valor 3 que é o caso em que enviar um canal em multicast custa 3 vezes mais do que enviar o canal em unicast, este é um valor exagerado, e que serve apenas para testar um caso extremo, os valores 1, 1.5 e 2 são os valores mais realistas.

Na figura 4.3, está a evolução da carga sobre a rede para os vários valores de β . Pode ver-se que o valor de β só tem influência, quando o número de utilizadores é grande o suficiente para um grande número de canais ser transmitido em multicast, quando isso acontece, β

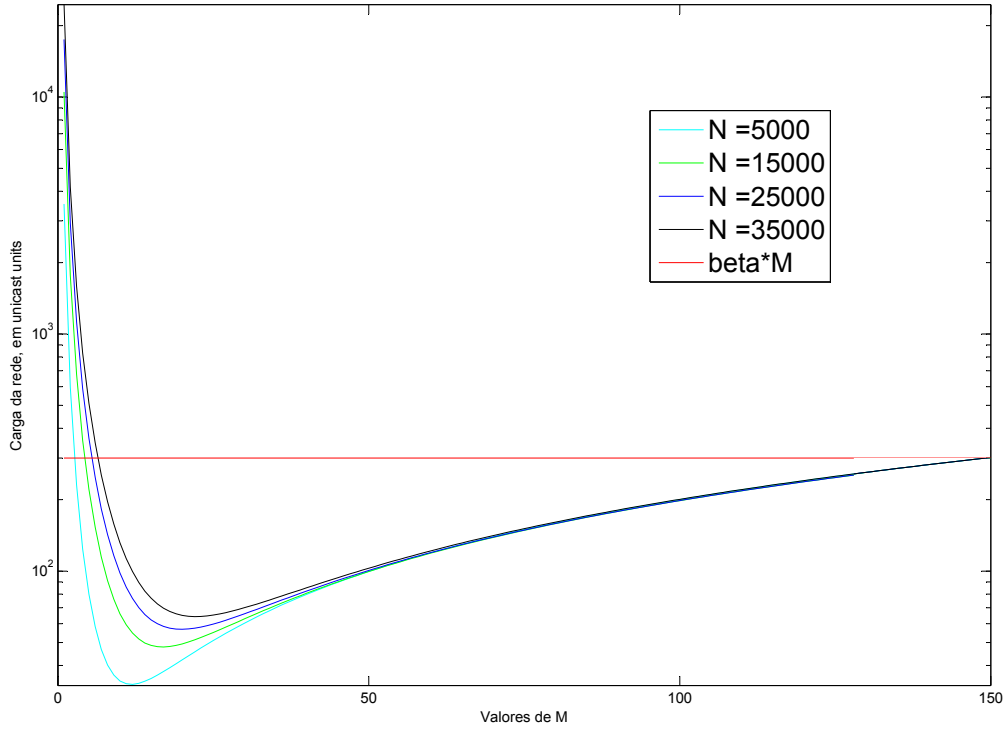


Figura 4.2: Evolução da carga no sistema dependendo do número de utilizadores activos

influencia o valor máximo de carga posta na rede, porque, quando todos os canais estão a ser transmitidos em multicast, a carga imposta à rede é βK , que é o caso em que todos os K canais são enviados em multicast.

4.2.2 Variação de α

O parâmetro α controla a distribuição da popularidade dos canais, assim, no cenário dinâmico α influencia o número de canais que vão ser transmitidos em unicast ou multicast. Uma distribuição de popularidade muito muito díspar, com um α muito elevado, em norma terá menos canais enviados em multicast do que uma distribuição de popularidade mais uniforme, porque os utilizadores vão estar mais concentrados nos canais mais populares, e os canais menos populares terão menos possibilidade de ter β utilizadores sintonizados. Na tabela 4.6 estão os parâmetros para os testes efectuados para testar a influência do parâmetro α . Neste caso os teste foram feitos com um número muito elevado de clientes, para mesmo com distribuições com α muito elevado, se ver que eventualmente a carga imposta sobre o sistema é, no máximo, $\beta * K$.

Na figura 4.4, está a evolução da carga sobre a rede para os vários valores de α . Pode ver-se que o valor de α influencia o número de utilizadores que o sistema precisa para que a carga máxima sobre a rede seja alcançada.

Utilizadores activos	Carga Mínima	M óptimo
5000	33	12
15000	48	17
25000	57	20
35000	64	22

Tabela 4.4: Valores M óptimos para diferentes números de utilizadores activos

	Cenário 1	Cenário 2	Cenário 3	Cenário 4
N	10000	10000	10000	10000
K	150	150	150	150
a	0.7	0.7	0.7	0.7
α	2.5	2.5	2.5	2.5
β	1	1.5	2	5

Tabela 4.5: Parâmetros de diferentes cenários para calcular o impacto de β na carga exercida sobre a rede

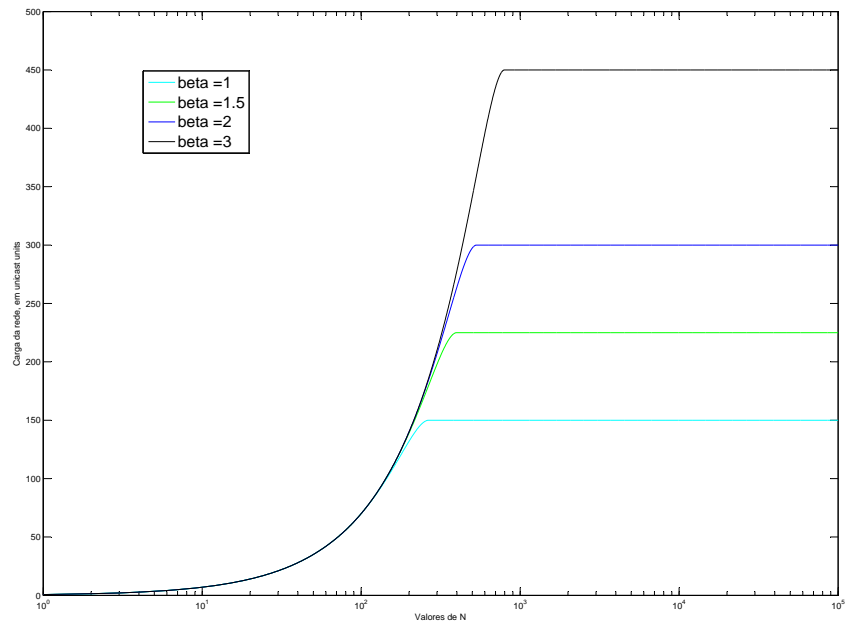


Figura 4.3: Evolução da carga no sistema dependendo do valor de β

	Cenário 1	Cenário 2	Cenário 3	Cenário 4
N	500000	500000	500000	500000
K	150	150	150	150
a	0.7	0.7	0.7	0.7
α	1.4	2	2.5	3
β	2	2	2	2

Tabela 4.6: Parâmetros dos cenários para iterativamente calcular o calor de M

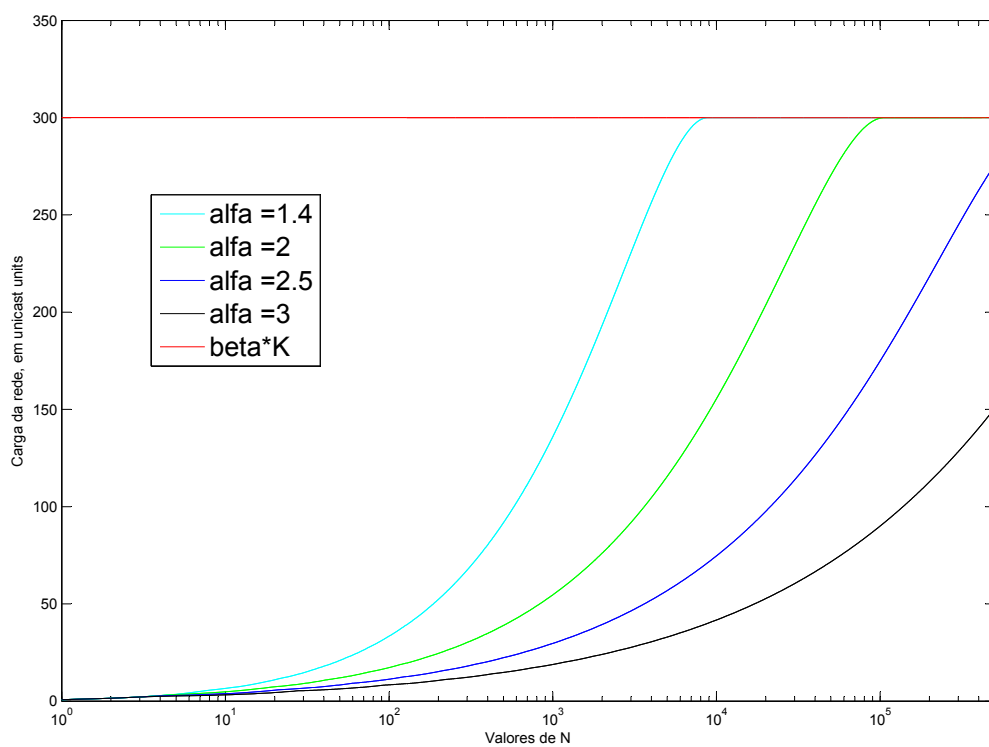


Figura 4.4: Evolução da carga no sistema dependendo do valor de α

Capítulo 5

Conclusões

A distribuição eficiente de canais IPTV é um problema grande para os operadores. É facilmente provado que enviar um canal em unicast para cada cliente é uma forma pouco eficiente de transmitir os sinais, porque a carga imposta na rede cresce muito rapidamente, e esgota a capacidade da rede rapidamente. A alternativa é enviar os canais em multicast, e criar sessões multicast a que os clientes aderem para ver os canais que querem. Mas mesmo assim, o custo de criar e manter sessões multicast pode não compensar enviar todos os canais sempre em multicast.

Para aferir a carga imposta na rede quer usando multicast, ou unicast, ou um modelo misto que combine unicast e multicast propomos um modelo, que simula quer o comportamento dos utilizadores, quer a popularidade dos canais, e que permite simular e constatar qual a carga a que a rede está sujeita em diferentes condições. A premissa para o nosso modelo era que, usando em conjunto, a transmissão em unicast e em multicast, poder-se-ia utilizar os recursos da rede de forma mais eficiente. No nosso caso a eficiência é a medida em que, em comparação com o que aconteceria usando apenas unicast ou multicast, a carga exercida na rede diminui ou aumenta.

Por isso, primeiro vamos definir, para o cenário só unicast, e para o cenário só multicast, para os mesmo parâmetros das simulações, quais seriam as cargas impostas na rede. Convém lembrar que em alguns gráficos apresentados no capítulo anterior estava uma linha que representava a equação $Y = \beta \cdot K$ que servia para proporcionar contexto em relação ao cenário de só multicast. Para um cenário só de unicast, e para um cenário só de multicast as cargas exercidas na rede são diferentes:

- Para um cenário apenas de unicast, para cada utilizador activo seria enviado um canal, mesmo que todos os utilizadores estivessem a ver o mesmo canal, por isso a procura de recursos do sistema seria $R_u = N$
- Para um cenário apenas de multicast, todos os canais seriam perpetuamente enviados em multicast, neste caso a procura de recursos do sistema seria $R_m = \beta \cdot K$.

Assim, vamos voltar a observar os gráficos do capítulo anterior, e observá-los sob um prisma diferente, ao invés de tentar otimizar os parâmetros do modelo, vamos observar a diferença entre os casos em que se transmitem canais só unicast, só multicast e o modelo misto.

5.1 Cenário Estático

A figura 5.1 mostra a carga no sistema para diferentes valores de α . Pode ver-se que o modelo misto, para qualquer valor de α é sempre melhor que o modelo que implica a transmissão de canais só em unicast. O modelo misto, para valores de α superiores a 1.5 é também superior a transmitir os canais apenas em multicast, as razões para este comportamento têm a haver com a distribuição dos utilizadores pelos canais como foi explicado no capítulo anterior.

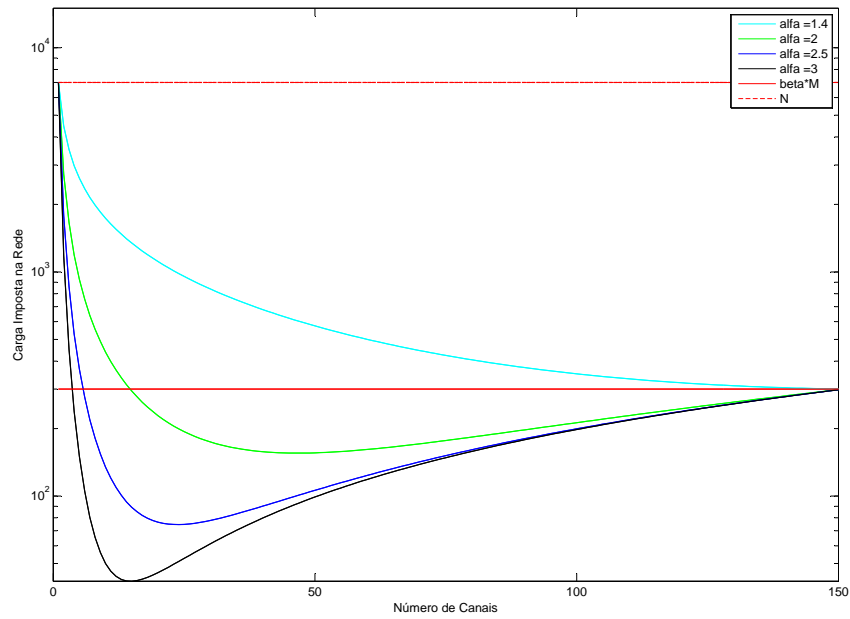


Figura 5.1: Comparação entre o modelo misto unicast/multicast e uso só de unicast ou multicast

Para diferentes valores de utilizadores activos, a carga do sistema é mostrada na figura 5.2. Este gráfico exemplifica ainda melhor os ganhos do uso de multicast, ou um sistema misto em relação a um sistema em que os canais são transmitidos apenas em unicast, porque conforme o número de utilizadores activos aumenta a diferença entre os canais enviados no caso de ser usado apenas unicast, e nos casos multicast e modelo misto é cada vez maior. Este comportamento deve-se ao facto de no pior caso o modelo misto ou o uso apenas de multicast exercerem uma carga na rede igual a βK , e o número K de canais ser em casos normais bastante inferior ao número N de canais que o uso de unicast impõe à rede.

5.2 Cenário Dinâmico

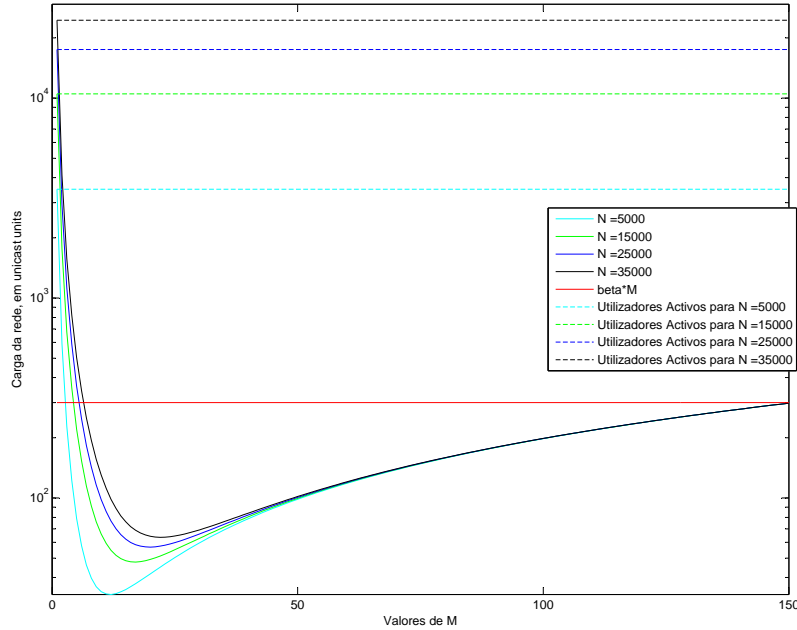


Figura 5.2: Comparação entre o modelo misto unicast/multicast e uso só de unicast ou multicast, para vários valores de utilizadores activos

No cenário dinâmico não é preciso analisar gráficos para tirar conclusões sobre a carga máxima a que o sistema pode estar sujeito, porque neste caso, a carga máxima que pode ser imposta ao sistema é $\beta * K$, que é o caso de todos os canais estarem a ser enviados em multicast. Podemos fazer esta afirmação porque no instante em que o custo de enviar o canal em multicast é igual ao custo de enviar o canal em unicast para um número limite β de utilizadores, o canal passa a ser enviado em multicast, havendo suficientes utilizadores activos no sistema, e dependendo também da distribuição de popularidade dos canais, todos os canais são enviados em multicast, e logo a carga máxima que é imposta na rede é $\beta * K$.

5.3 Trabalho Futuro

O trabalho futuro deste dissertação inclui, obrigatoriamente, o teste que em simuladores, quer em equipamento dos métodos propostos para suportar multicast usando Carrier Ethernet. Este teste não foi possível nesta dissertação porque, durante os trabalhos referentes a esta dissertação os protocolos IEEE802.Qay e MPLS-TP não estavam totalmente definidos. Assim foi apenas feita uma análise ao modelo, análise esta, que pode, ao ser passada para um cenário real provar-se errada, porque dependendo dos protocolos pode até provar-se que o custo de enviar um canal em multicast é igual ao custo de enviar um canal em unicast, ou até pode ser estabelecido um limiar (um certo número de utilizadores sintonizados num canal) acima do qual em vez de enviar o canal em multicast é mais eficiente fazer broadcast do mesmo.

Apêndice A

Ethernet

A tecnologia Ethernet como já foi dito foi desenvolvida nos anos 70 com o intuito de ligar computadores numa rede local, e está definida na norma 802.3 do IEEE. Funcionalmente a tecnologia Ethernet divide as camadas física e de ligação do modelo OSI em três sub-camadas

- *Subcamada física* especifica a interface física e a cablagem, é responsável pela transmissão e recepção dos sinais da rede. Tipicamente aparece sob a forma de uma placa de rede (NIC - Network Interface Card), que se liga ao terminal que requer conectividade Ethernet. Cada NIC tem um endereço único de 6 bytes, que identifica univocamente o terminal a que está ligado, este endereço é conhecido como o endereço MAC.
- *Subcamada de controlo de acesso do meio* define funções independentes do meio físico de transmissão e que dizem respeito à forma lógica como a informação é enviada e recebida, a detecção de erros e colisões. Tem duas funções principais:
 - **Encapsulamento da informação** esta subcamada é responsável pela construção das *frames* Ethernet a enviar e é responsável por detectar erros em *frames* Ethernet recebidas
 - **Controlo de acesso ao meio** esta sub camada é responsável evitar colisões no envio de *frames*, e por lidar com o meio de transmissão partilhado.
- *Camada de controlo de ligações lógicas* esta subcamada é responsável por pela comunicação entre a camada MAC e as camadas superiores, ocupa-se principalmente multiplexação das *frames* transportadas pela camada MAC e fornece mecanismos de controlo de fluxo, aviso e recuperação.

A forma como as camadas Ethernet se ligam com as camadas OSI está explicita na figura A:

A camada física da Ethernet corresponde à camada física do modelo OSI e, mas a camada de ligação de dados do modelo OSI corresponde às subcamadas de controlo de acesso ao meio e de controlo de ligações lógicas da Ethernet.

A.1 Elementos da Rede

Os elementos das redes que usam Ethernet dividem-se em três grupos:

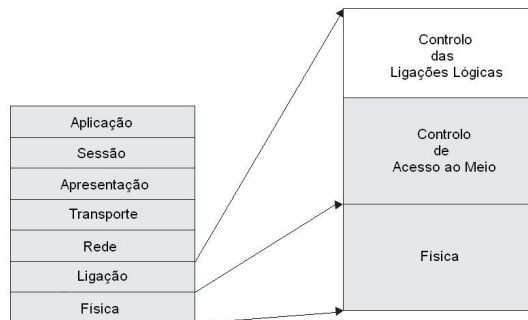


Figura A.1: camadas OSI

- **Equipamentos terminais** dispositivos que são a fonte ou o destino da informação
- **Equipamentos de comunicação** dispositivos que servem para interligar dispositivos terminais, ou troços da rede, encaminhando mensagens Ethernet com base no seu endereço de destino.
- **Equipamento de ligação** meio físico que liga os componentes da rede e sobre o qual viajam os dados. Tipicamente é um cabo UTP (unshielded twisted pair), mas também pode ser um cabo de fibra óptica ou até o ar no caso de wireless LANs.

A.1.1 Ethernet Switches e Bridges

Existem dispositivos que operando ao nível da camada 2 do modelo OSI permitem ligar 2 ou mais segmentos de LAN, permitindo assim criar redes maiores. Tais dispositivos chamam-se Switches ou Bridges, a diferença entre os dois é que o switch funciona com hardware, e a bridge funciona por software.

Estes dispositivos têm várias portas Ethernet (cada uma ligada a uma sub-rede diferente) e funcionam aprendendo os endereços MAC alcançáveis por cada uma das suas portas, e construindo uma tabela de encaminhamento com essa informação, assim quando uma *frame* chega à bridge esta ou a descarta ou a encaminha pela porta que oferece conectividade com o dispositivo de destino, e só por essa porta. Caso seja uma *frame* com endereço de broadcast, é encaminhada por todas as portas, menos pela porta por onde chegou.

A.2 Ethernet *frames*



Figura A.2: *frame* Ethernet

A tecnologia Ethernet tem um tipo de mensagem definido, mostrado na figura A.2, o tamanho máximo de uma *frame* Ethernet é 1526 bytes que corresponde a 26 bytes do cabeçalho

e outros dados de controlo e a um payload máximo de 1500 bytes. Devido a inovações como V-LANs e MPLS o cabeçalho de mensagens Ethernet foi modificado para incluir mais campos. Também foi criada uma classe de *frames* chamada de *jumbo frames* que podem ter mais de 9000 bytes, e destinam-se a soluções comerciais. Na figura A.2 está uma *frame* Ethernet com os respectivos campos realçados:

Tabela A.1: Campos de uma *frame* Ethernet

Campo	Bytes	Descrição
Preâmbulo	7	Indica que uma <i>frame</i> está prestes a ser difundida
SOF (Start of Frame)	1	Indica o início da frame
Endereço de destino (DA)	6	Identifica o terminal destinatário da <i>frame</i>
Endereço de origem (SA)	6	Identifica o terminal de origem da <i>frame</i>
Tamanho/Tipo	2	Identifica o tamanho o número de bytes enviados ou o tipo de <i>frame</i>
Dados/ <i>Payload</i>	46-1500	Dados transportados
FCS (Frame Check Sequence)	4	Valor usado para garantir a integridade da <i>frame</i>

A.3 Funcionamento Ethernet

A Ethernet tem um principio de funcionamento simples:

1. O terminal que quer transmitir verifica se o meio de transmissão está ocupado
 - Se estiver ocupado espera até poder transmitir
 - Se estiver livre o terminal transmite uma *frame*
2. volta a 1. até terminar de mandar todas *frames*

Este algoritmo é conhecido por *Carrier Sense Multiple Access with Collision Detection*, ou CSMA/CD, e foi desenvolvido para evitar e detectar colisões entre *frames*. Este algoritmo ainda é usado em ligações Ethernet a 10Mbps e FastEthernet a 100Mbps, mas foi abandonado nas ligações Gigabit Ethernet a 1Gbps. Quando uma *frame* é recebida por um terminal, este verifica se o endereço de destino da *frame* é o seu endereço MAC, se for o terminal processa a *frame*, caso não seja a *frame* é descartada. Há também o caso particular de transmissão em broadcast, neste caso o DA está preenchido apenas com 1s e todos os terminais que recebem a *frame* processam-na.

No funcionamento das redes Ethernet surge um problema quando os dispositivos são interligados de forma a que existam ciclos, ou "loops" entre eles, neste caso duas características da tecnologia Ethernet conjugam-se para criar um enorme problema:

- Os dispositivos Ethernet quando recebem uma *frame* multicast ou broadcast enviam-na por todas as portas, menos pela porta por onde a receberam.

- As *frames* Ethernet não têm um campo com o funcionamento do campo "Time to Live" dos pacotes IP.

Com estas duas características da tecnologia Ethernet, basta apenas uma *frame* multicast ou broadcast para que aconteça um fenómeno chamado "broadcast storm". Este fenómeno acontece quando um dispositivo recebe uma *frame* broadcast de um terminal, de seguida envia essa *frame* broadcast por todas as suas portas, menos pela porta por onde recebeu a *frames*, e todos os dispositivos que estão directamente ligados a esta bridge quando recebem esta mensagem têm o mesmo comportamento, ou seja enviam a *frame* por todas as portas, menos pela porta por onde a *frame* chegou. O problema acontece quando existe um ciclo na rede, e a bridge que enviou a *frame* broadcast recebe uma cópia dessa mesma *frame*, e a primeira bridge volta a re-enviar a *frame*, assim como todos os dispositivos que recebem uma segunda cópia da *frame*, este comportamento faz com que haja várias cópias da *frame* a circular na rede, e sempre que um dispositivo recebe uma cópia da *frame* volta a re-enviá-la por todas as portas menos pela porta onde recebeu a *frame*, originando ainda mais cópias da *frame*. Caso as *frames* Ethernet tivessem um campo "Time to Live", eventualmente as cópias seriam descartadas, mas como as *frames* Ethernet não têm esse campo as *frames* broadcast são re-enviadas até eventualmente ocuparem quase toda a largura de banda disponível e degradarem o desempenho da rede. Há várias soluções para este problema, podia alterar-se o comportamento dos dispositivos Ethernet aquando de uma *frame* broadcast ou multicast, podia alterar-se a *frame* Ethernet para incluir um campo "Time to Live", podiam rejeitar-se as topologias de rede que incluíssem ciclos ou podia introduzir-se um protocolo que mesmo que uma determinada topologia incluísse ciclos, desligasse, por software portas de bridges de modo a eliminar os ciclos, e assim o problema, e foi esta última opção a adoptada, surgindo o protocolo IEEE802.1D *Spanning Tree Protocol* (STP).

O protocolo *Spanning Tree* divide a topologia da rede em duas partes, a topologia física e a topologia activa [15]. A topologia física é constituída por todos os dispositivos e todas as ligações físicas, a topologia activa é constituída por todos os dispositivos e por um conjunto de ligações activas, o conjunto das ligações activas é um subconjunto das ligações físicas que consiste nas ligações que não são desligadas para prevenir a ocorrência de ciclos.

A.4 IEEE 802.1Q - Virtual LAN

O primeiro esforço para melhorar a tecnologia Ethernet, foi a norma 802.1Q - Virtual LAN. Esta tecnologia permite criar várias LANs virtuais sobre a mesma infraestruturas, isto é feito inserindo uma *tag* de 12 bits chamada de Q-TAG, VLAN-TAG ou VLAN-ID no cabeçalho Ethernet que identifica a VLAN da *frame* [12]. Assim esta tecnologia permite gerir as VLAN de forma independente, e atribuir a cada VLAN um papel diferente dentro da mesma organização. Cada VLAN tem um domínio de *broadcast* independente, e um conjunto de portas atribuído [10],

Esta tecnologia funciona bem dentro das fronteiras de uma organização, separando o tráfego entre as diferentes VLANs, e aumentando a resistência a falhas visto que usando o algoritmo MSTP em que cada VLAN tem a sua própria *Spanning Tree*. Mas quando se tenta aplicar esta tecnologia às redes de transporte rapidamente surgem problemas quanto a essa utilização:

- Dificuldade em atribuir os VLAN-IDs. Tradicionalmente quem distribui os VLAN-IDs

é o gestor da rede privada, mas no caso das redes de transporte quem atribui os VLAN-IDs é o prestador de serviços, que neste caso teria de certificar-se que os VLAN-IDs que atribui são únicos quer na rede de transporte quer nas redes privadas onde os gestores de rede querem manter o controlo sobre a atribuição de VLAN-IDs.

- Problema de escalabilidade, visto que há um número limitado, a *tag* tem 12 bits o que limita o número de VLAN-IDs utilizáveis a 4094 visto que 2 estão reservados para administração. Este facto implica que cada prestador de serviços estaria limitado a 4094 clientes. O que é claramente insuficiente [22] [14] [10].
- O uso de VLAN herda também o problema de escalabilidade da tecnologia Ethernet que é o uso de um esquema de endereçamento plano, que implica que cada dispositivo conheça o endereço MAC do dispositivo na rede para o qual quer enviar uma *frame*, aplicar a tecnologia VLAN às redes de transporte iria implicar que todos os *switches* da rede tivessem tabelas de encaminhamento enormes com todos endereços quer da rede de transporte, quer das redes privadas que pertencessem à mesma VLAN.

VLAN Frame

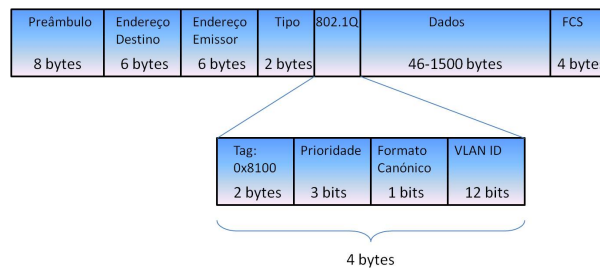


Figura A.3: Frame VLAN

Na figura A.3 está uma *frame* VLAN, foi adicionada uma *tag*, que se divide em 4 campos. Os mais importantes são:

- VLAN-ID: É um número que identifica a VLAN a que a *frame* pertence.
- Prioridade: É um campo que possibilita 8 níveis de prioridade ordenados do valor com mais prioridade para o valor com menos prioridade: 7-6-5-4-3-0-2-1 [16] (pag 383). O valor de prioridade é usado para o *forwarding* das frames, porque apesar de tipicamente cada porta da Bridge estar atribuída a uma VLAN, podem haver portas onde este campo indica quais a *frames* que devem ter prioridade.

Apêndice B

SDH

B.1 SDH

A tecnologia SDH(Synchronous Digital Hierarchy) pertence à primeira geração de redes ópticas, tal como a Ethernet funciona como camada 2 do modelo OSI, fornecendo serviços de construção de tramas, multiplexagem dos dados enviados sobre o meio físico, controlo de erros e sub-camada MAC para protocolos de nível superior que a usam como camada de ligação de dados e camada física. A camada SDH fornece:

- Ligações do tipo comutação de serviços extremo-a-extremo da rede
- Multiplexação eficiente das ligações de baixo ritmo para obter ligação de débito mais elevado
- Um método eficiente, usando ponteiros, para nos nós intermédios extrair fluxos de ritmos mais baixos, a partir de fluxos mais elevados
- Um elevado grau de fiabilidade entre 99,99% e 99,999%
- Capacidades OAM, cerca de 5% do tráfego é reservado para gestão da rede
- Uma tecnologia normalizada de transporte a altas velocidades que permitem a satisfação de SLAs acordados com clientes

O bloco fundamental do SDH é o STS-1 (Synchronous Transporte Signal-1), logicamente pode ser representado por uma estrutura com 90 colunas e 9 linhas, contendo cada célula da estrutura um byte, onde as 3 primeiras colunas são reservadas para *overhead* de transporte, e as restantes 87 colunas são usadas para transportar a *payload* STS, esta estrutura é transmitida a cada $125\mu s$, resultando numa velocidade de 51.840 Mbps.

O SPE (Synchronous Payload Envelope) é uma estrutura de 87x9 bytes que é transportada na *payload* do STS, os limites do SPE não têm obrigatoriamente de estar alinhados com os limites do STS, podendo um SPE estar distribuído por mais que um STS. Para o SDH saber onde está o início de cada SPE dentro de cada STS, cada SPE tem um *overhead* próprio, que inclui um byte que serve de ponteiro para o início do SPE dentro do STS. Quando há variações na velocidade de relógio entre sinal SPE e a rede, apenas o ponteiro é alterado para reflectir essa mudança. O SDH consegue também multiplexar sinais de ritmos mais baixos para atingir ritmos de transmissão mais elevados. Os ritmos de transmissão SONET e SDH estão representados na tabela B.1

Tabela B.1: Ritmos de transmissão SDH e SONET

Sinal SONET	Velocidade (Mbps)	Sinal SDH
STS-1, OC-1	51,840	STM-0
STS-3, OC-3	155,520	STM-1
STS-12, OC-12	622,080	STM-4
STS-48, OC-48	2,488,320	STM-16
STS-192, OC-192	9,953,280	STM-64
STS-768, OC-768	39,813,120	STM-256

B.1.1 Elementos da Rede

As redes SDH usam vários componentes para inserir e retirar eficiente fluxos de tráfego na rede, para regenerar o sinal óptico e para comutar o tráfego. Assim os componentes SDH dividem-se em:

- Terminais de linha que multiplexam e demultiplexam fluxos de tráfego
- Add-Drop Multiplexers que fornecem uma forma eficiente para extrair a parte do tráfego destinado a uma determinado nó, permitindo a passagem do restante tráfego
- Digital Cross Connects usados em redes maiores para comutar um grande número de fluxos de tráfego
- Regeneradores para regenerar o sinal sempre que necessário

B.1.2 SDH *frames*

A figura B.1 mostra duas *frames* SONET que transportam um SPE STS-1 que está distribuído por ambas. Como referido anteriormente cada *frame* SONET é uma matriz de 90x9 bytes transmitida a cada 125 μ s. Na *frame* onde começa o SPE estão também indicados os bytes indicados para *overhead* de secção de regenerador e do multiplexer, e também os bytes reservados para administração onde se encontra o ponteiro para o início do STS.

B.1.3 Funcionamento SDH

As redes SONET/SDH são implementadas de três formas:

- Configuração linear
- Configuração ponto-a-ponto
- Configuração em anel (baseado em *token-ring*)
 - UPSR (Unidirectional Path-Switched Ring) anel que usa 2 fibras que funcionam em sentido contrário, uma de trabalho, e uma de segurança que é activada quando

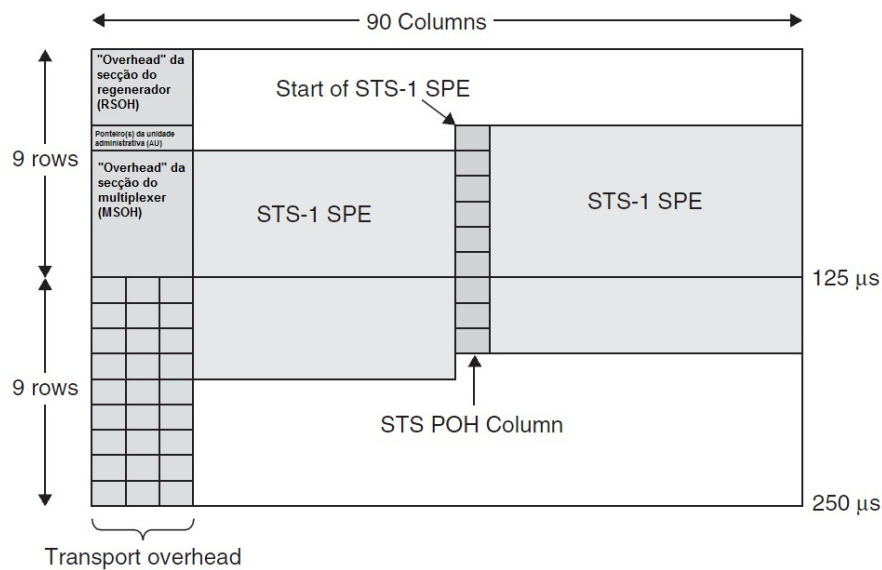


Figura B.1: STS distribuido por duas frames SONET, imagem disponível em [21]

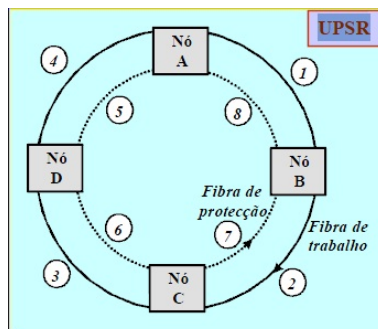


Figura B.2: Exemplo de anel UPSR, imagem disponível em [21]

se detecta uma falha na fibra de trabalho, este tipo de protecção é chamado de 1+1. Na figura B.2 encontra-se um exemplo de anel UPSR, imagem disponível em [21].

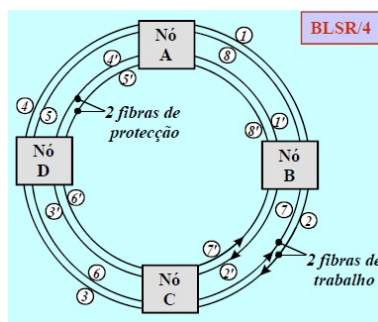


Figura B.3: Exemplo de anel BLSR/4, imagem disponível em [21]

- BLSR (Bidirectional Line-Switched Ring), também conhecidos por BHSR (Bidirectional Self-Healing Ring) há duas variantes deste tipo de anel: BLSR/2 e BLSR/4. A diferença entre os dois é o número de fibras usadas, 2 e 4 respectivamente. A maior diferença deste tipo de anéis para os anéis UPSR é que neste tipo de anel, o tráfego é transportado nas duas direcções simultâneamente. Na figura B.3 está exemplificado um anel BLSR4. O uso de 4 fibras faz com que seja possível, em caso de falha, implementar comutação de vão no caso de ocorrer uma falha num emissor ou receptor, ou implementar comutação de anel quando é cortado o cabo que contenha as quatro fibras do anel.

O anel BLSR/2 funciona de modo indêntico ao BLSR/4 mas usa duas fibras, cada uma funciona numa direcção, e cada uma guarda metade da capacidade para encaminhar o tráfego em caso de falha.

Apêndice C

C - Métodos probabilísticos

C.1 Variável de Bernoulli

Se A for um acontecimento que resulte de uma experiência aleatória com apenas dois resultados possíveis: sucesso ou insucesso. Então A é uma variável de Bernoulli com $P(\text{sucesso}) = p$ e $P(\text{insucesso}) = (1 - p)$ [26].

C.2 Distribuição Binomial

Uma distribuição binomial, é uma distribuição que respeita 3 regras:

1. Em cada ensaio há apenas dois resultados possíveis: sucesso ou insucesso
2. Os ensaios são independentes entre si
3. Cada ensaio tem $P(\text{sucesso}) = p$ e $P(\text{insucesso}) = (1 - p)$

É fácil de ver, através de 1 e 3 que cada ensaio pode ser encarado como uma variável de Bernoulli. Assim uma distribuição Binomial permite calcular a probabilidade de após n ensaios terem havido k sucessos. A função de densidade de probabilidade é definida por:

$$Pr[X = n] = C_n^k (1 - p)^{n-k} p^k = \frac{n!}{(n - k)!k!} (1 - p)^{n-k} p^k$$

Esta função é fácil de deduzir, pela independência dos acontecimentos, a probabilidade de de haverem k sucessos, em qualquer ordem é dada por C_n^k , e novamente pela independência dos acontecimentos a probabilidade de haver k sucessos é p^k . Como só há dois resultados possíveis, a probabilidade de insucesso é complementar da probabilidade de sucesso, e se houve n ensaios e k sucessos então houve $n - k$ insucessos, e a probabilidade de haver $n - k$ insucessos é $(1 - p)^{n-k}$

C.3 Distribuição Multinomial

A distribuição Multinomial é uma generalização da distribuição Binomial, mas em vez de os acontecimentos terem apenas dois resultados possíveis, o resultado de cada ensaio sai de

um conjunto finito com k possibilidades. Uma distribuição multinomial é uma distribuição com as seguintes propriedades [?]:

- Consiste em n ensaios consecutivos
- Cada ensaio tem um número finito de resultados possíveis
- A probabilidade de sucesso de cada resultado é constante em todos os ensaios
- Os ensaios são independentes

A função de densidade de probabilidade é dada por:

$$Pr = [c_1 = b_1, \dots, c_K = b_K] = \frac{n!}{b_1! \dots b_K!} \pi_1^{b_1} \dots \pi_K^{b_K}$$

C.4 Distribuição de Zipf

A distribuição de Zipf, é uma distribuição, descoberta de forma empírica por George Zipf nos anos 30, para tentar prever a frequência que as palavras apareciam em livros. A distribuição de Zipf depende de 3 parâmetros:

- O número dos elementos
- O classificação dos elementos
- O expoente que caracteriza a distribuição

C.4.1 Implicação do parâmetro β

A distribuição de Zipf que usamos na dissertação é:

$$\pi_k = dk^{-\alpha}, \text{ para } k = 1, 2, \dots, K \quad (\text{C.1})$$

Neste caso, o número de elementos é o número de canais, K , a classificação dos elementos é k , e α é o expoente que caracteriza a distribuição. A figura C.1 tem imagens de diferentes distribuições de popularidade, caracterizadas por diferentes valores de α .

Pode constatar-se que um valor α de 0 corresponde a uma distribuição uniforme, em que todos os canais têm igual probabilidade de serem vistos. E que conforme o valor de α aumenta, a distribuição de popularidade altera-se e um número restrito de canais começa a sobressair como os mais populares.

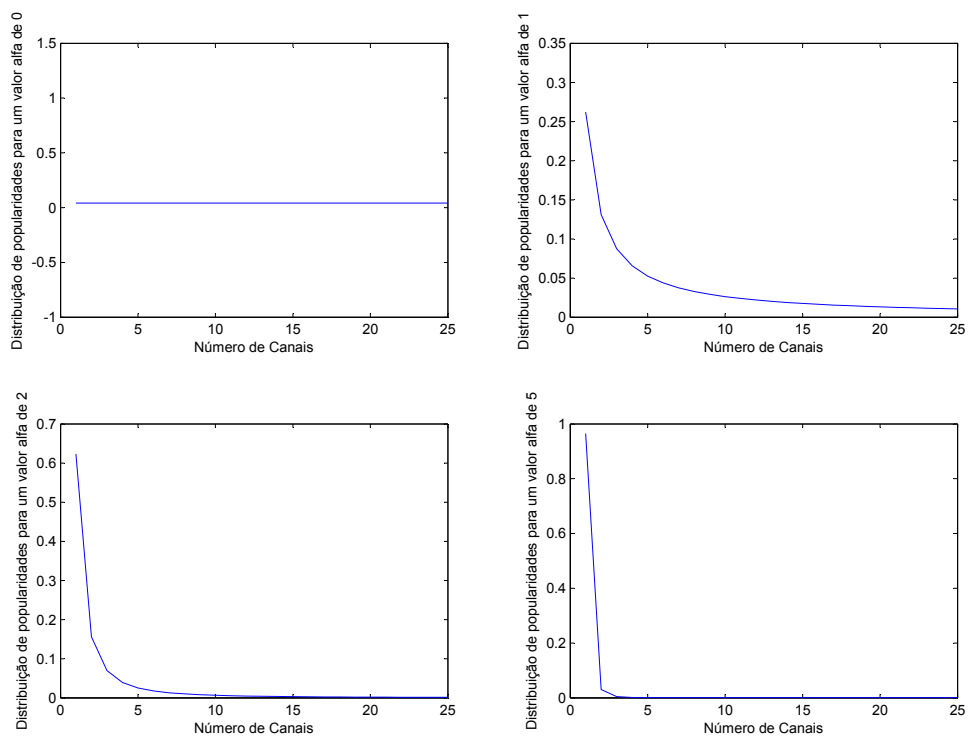


Figura C.1: Vectors de distribuição de popularidade para diferentes valores de α

Apêndice D

Multi-Protocol Label Switching - MPLS

O protocolo MPLS, definido na RFC3031 [23] opera entre a camada de ligação de dados e a camada de rede, a forma como o protocolo encaixa no modelo OSI está na figura D.1. Este protocolo usa encaminhamento baseado em *labels*, cada dispositivo tem uma lista de *labels* e informação sobre como encaminhar pacotes marcados com cada uma das *labels*.

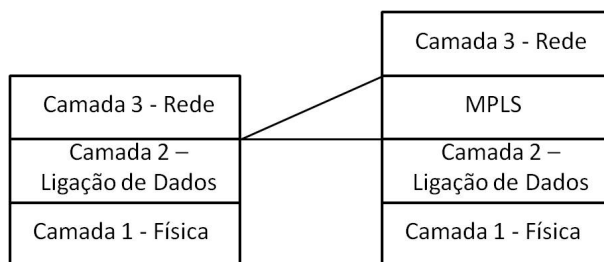


Figura D.1: Enquadramento do protocolo MPLS no modelo OSI

Os caminhos que os pacotes tomam dentro da rede MPLS são chamados *Label Switched Paths*, ou LSPs, estes LSPs são identificados por *labels* [8] e podem ser estabelecidas quer de forma manual, quer usando um protocolo de sinalização. Este protocolo usa um formato de pacote próprio onde encapsula os pacotes de camadas superiores. O cabeçalho desse pacote está na figura D.2. Este cabeçalho pode variar de tamanho, porque várias *labels* podem ser empilhadas no mesmo cabeçalho. RFC

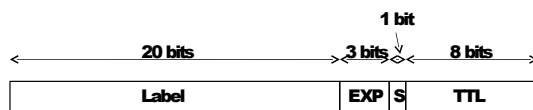


Figura D.2: Cabeçalho usado pelo protocolo MPLS

O cabeçalho tem 4 campos, o campo *Label* com 20 bits que guarda a *label* com que o pacote é encaminhado, o campo *EXP* (*Experimental*) com 3 bits, que aquando da definição

do protocolo MPLS ficou reservado para uso futuro, embora já seja usado há bastante tempo para implementar QoS, o campo *S (End of Stack)* com 1 bit que sinaliza se a pilha de *labels* está vazia ou não, e o campo *TTL (Time to Live)* com 8 bits que tem a mesma função que o campo homônimo no cabeçalho dos campos IP que é o de impedir que os pacotes circulem eternamente na rede. Este cabeçalho fica colocado entre o cabeçalho de camada 2 e o cabeçalho de camada 3, o MPLS como o nome indica é independente quer da tecnologia de camada 2 quer da tecnologia de camada 3 usadas, no entanto enquanto a tecnologia de camada 2 pode variar muito: Ethernet, SDH, Frame Relay e a tecnologia de camada 3 utilizada em conjunto com o protocolo MPLS geralmente é o IP este protocolo é também conhecido como IP/MPLS. Na figura D.3 está um exemplo de um pacote MPLS com um pacote IP no *payload*.

O protocolo MPLS introduz também o conceito de "Forwarding Equivalent Class", ou FEC, uma FEC é definida como o conjunto de todos os pacotes que são encaminhados pelo mesmo caminho, ou conjunto de caminhos, dentro da rede [23]. Para além de serem encaminhados pelas mesmas rotas através da rede, as FEC podem ser definidas por mais atributos, como a rede de origem dos pacotes, ou a CoS. Os pacotes são classificados em diferentes FECs no LER quando são encapsulados em pacotes MPLS.



Figura D.3: Pacote MPLS com cabeçalho e *payload*

D.0.2 Equipamentos

O protocolo MPLS divide os equipamentos em duas classes, com base nas funções que cada equipamento tem:

- Label Edge Routers (LER): Estes routers estão localizados na fronteira das redes e são responsáveis por encapsular a informação em pacotes MPLS e encaminhá-los de acordo com a sua tabela de encaminhamento. No sentido contrário, quando recebem um pacote MPLS encaminhado para fora da rede são responsáveis por retirar o cabeçalho MPLS e encaminhá-lo. Quando o LSR encapsula um pacote IP, pode copiar o valor do campo TTL do pacote IP para o campo TTL do pacote MPLS. O protocolo MPLS tem uma opção chamada *penultimate hop popping*, que foi criada para evitar que os LERs sejam sobrecarregados com pacotes para retirar o cabeçalho MPLS, com esta opção, o último LSR antes de enviar o pacote para o LER retira o cabeçalho MPLS e envia o pacote como um pacote IP normal.
- Label Switching Routers (LSR): Estes routers estão localizados no interior da rede, e encaminham os pacotes MPLS de acordo com a sua própria tabela de encaminhamento e a *label* que esteja no topo da pilha. Quando enviam um pacote MPLS, estes dispositivos decrementam o valor do campo TTL do pacote MPLS, e caso este valor atinja o valor 0, o pacote é descartado. Esta tabela de encaminhamento contém *labels*, e não endereços IP, e o modo como é preenchida é explicado mais à frente.

Estas definições são aplicadas aos routers para cada LSP, pode acontecer que para um LSP um determinado router actue como um LER e para outro LSP actue como LSR, depende da topologia da rede, e das rotas percorridas por cada LSP. No geral, para além da

encapsulamento ou desencapsulamento de pacotes IP em pacotes MPLS, cada router MPLS pode efectuar 3 operações sobre as *labels* de cada pacote:

- *Label Swap*: Os routers podem substituir a *label* do topo na pilha, por outra. Todos os routers ao longo de um LSP efectuem esta operação.
- *Label Push*: Os routers podem adicionar *labels* ao cabeçalho MPLS, as *labels* são adicionadas em pilha, ou seja a última a ser adicionada é a próxima a ser retirada. Quando os pacotes são encaminhados para uma rede MPLS e lhes é adicionado um cabeçalho MPLS, este cabeçalho já contém uma *label*, no entanto enquanto percorrem a rede, os LSRs podem adicionar mais *labels*. *Labels* adicionais podem ser adicionadas após a troca da *label* do topo na pilha.
- *Label Pop*: Os routers podem retirar a *label* que está no topo da pilha. Esta operação pode ser efectuada para retirar *labels* que estejam empilhadas, ou pode indicar ao router que deve retirar todo o cabeçalho MPLS, esta indicação é feita através do bit *EOS*, quando este bit é "1" indica que não há mais *labels* na pilha.

Na figura D.4 está o exemplo de uma rede MPLS com 5 LSRs, 5 LERs e 3 LSPs. Neste exemplo é possível ver que o router LER4 tem diferentes papéis para diferentes LSPs, para o LSP representado a Vermelho actua como LER porque recebe os pacotes fora da rede MPLS e é responsável pelo encapsulamento e encaminhamento desses pacotes para dentro da rede MPLS, mas para o LSP representado a Verde o router LER4 actua como um LSR mudando a *label* e encaminhando o pacote para o router LER2. Também é possível ver neste exemplo que o router LSR1 adiciona uma *label* aos pacotes pertencentes ao LSP 2, mas antes de adicionar uma *label* nova ao topo da pilha troca *label* que fica na segunda posição da pilha. Na imagem está também um exemplo de agregação de LSPs, em que no router LSR 5 o LSP 2 e o LSP 3 são agregados em apenas um LSP e encaminhados para o router LER4.

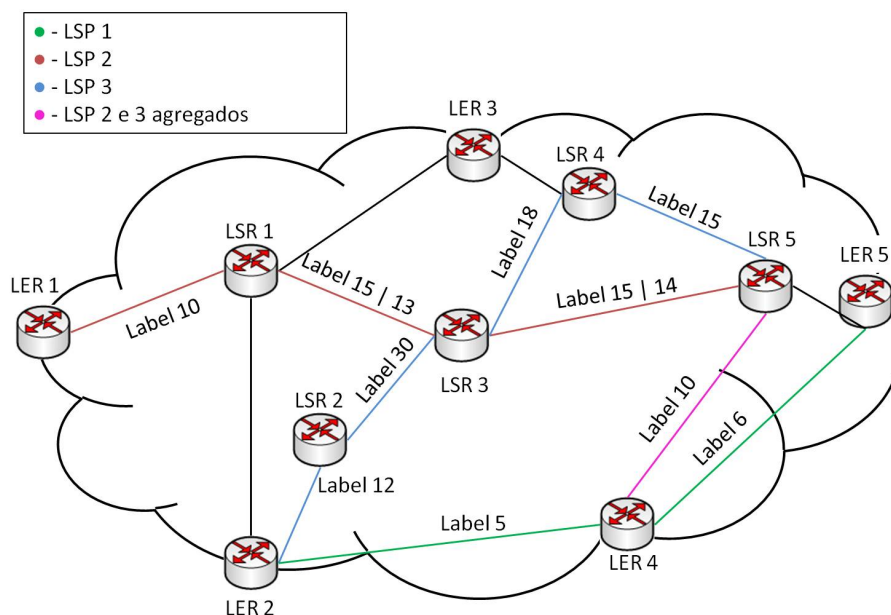


Figura D.4: Exemplo de uma rede que usa o protocolo MPLS, com 2 LSPs definidos

D.1 Estabelecimento de Ligações e *multicast*

O protocolo MPLS pode estabelecer ligações de várias formas, as formas mais comuns são as ligações serem criadas de forma manual aquando da configuração dos routers, ou usando um protocolo de sinalização tipicamente o protocolo *Label Distribution Protocol*, ou protocolo *Resource Reservation Protocol - Traffic Engineering*.

- **Configuração Manual:** O operador pode escolher configurar manualmente todos os LSRs da rede para fazerem o encaminhamento dos pacotes. Esta solução pode ser demorar muito tempo, e é muito susceptível quer a falhas de configuração, por erro humano podem ser configurados LSPs com ciclos, ou LSPs que não encaminham o tráfego para o LSR correcto, quer a falhas da rede, porque se há um problema com um router, ou uma ligação o comportamento da rede não se adapta a essa realidade e pode continuar a encaminhar tráfego pelos dispositivos ou ligações que não estão a funcionar.
- ***Label Distribution Protocol (LDP)*:** Este protocolo definido na RFC3036 [2], é um protocolo que permite aos LSRs estabelecerem sessões para anunciarem *labels*. Dois routers MPLS adjacentes tomam conhecimento um do outro através de mensagens "Hello" multicast enviadas para a porta UDP 646. Este protocolo tem de funcionar em conjunto com um protocolo de *routing* que permita aos routers tomarem conhecimento das rotas entre eles. Após os LSRs tomarem conhecimento das rotas entre eles, os LSPs criados seguem o caminho mais curto de acordo com o protocolo de encaminhamento usado.
- ***Resource Reservation Protocol - Traffic Engineering (RSVP-TE)*:** Este protocolo definido na RFC3209 [4], é um protocolo capaz de fazer reservas de largura de banda e definir parâmetros de QoS para os LSPs configurados na rede. Este protocolo também é dependente de um protocolo de encaminhamento para descoberta e cálculo dos caminhos mais curtos entre os LSPs.

Bibliografia

- [1] L. Andersson. "EXP field" renamed to "Traffic Class field". RFC 5462 (Standards Track), 2009.
- [2] L. Andersson, Nortel Networks Inc., P. Doolan, Ennovate Networks, N. Feldman, IBM Corp, A. Fredette, PhotonEx Corp, B. Thomas, and Inc Cisco Systems. "LDP Specification ". RFC 2036 (Standards Track), 2001.
- [3] Zlatka Avramova, Danny De Vleeschauwer, Sabine Wittevrongel, and Herwig Bruneel. Capacity Gain of Mixed Multicast/Unicast Transport Schemes in a TV Distribution Network. *IEEE TRANSACTIONS ON MULTIMEDIA*, 11(5):918–931, AUG 2009.
- [4] D. Awduche, Inc. Movaz Networks, L. Berger, D. Gan, Inc. Juniper Networks, T. Li, Inc. Procket Networks, V. Srinivasan, Inc. Cosine Communications, G. Swallow, and Inc. Cisco Systems. "RSVP-TE: Extensions to RSVP for LSP Tunnels". RFC 3209 (Standards Track), 2001.
- [5] Nigel Bragg. Ethernet as an infrastructure; pbb-te and beyond, 2007.
- [6] Brocade. Leveraging the benefits of provider backbone bridges. url, 2009.
- [7] Roberta Crescentini e Fabio Valente Christian Addeo, Giorgio Cazzaniga. On qos mechanism profiling in mpls-tp transport networks. *Bell Labs Technical Journal*, 14(4):227–241, 2010.
- [8] Y Donoso. *Network Design for IP Convergence*. Auerbach Publications,, Boca Raton, Florida, EUA, 2009.
- [9] Samer Salam e Ali Sajassi. Provider backbone bridging and mpls: Complementary technologies for next-generation carrier ethernet transport. *IEEE Communications Magazine*, 22(1):77–83, 2008.
- [10] Don Fedyk e David Allan. Ethernet data plane evolution for provider networks. *IEEE Communications Magazine*, 22(1):84–89, 2008.
- [11] Paul Bottorff e Panagiotis Saltsidis. Scaling provider ethernet. *IEEE Communications Magazine*, 22(1):104–108, 2008.
- [12] Pasula Reddy e Sam Lisle. Ethernet aggregation and transport infrastructure oam and protection issues. *IEEE Communications Magazine*, 22(1):152–159, 2009.
- [13] D. Frost, S. Bryant, Cisco Systems, M. Bocci, and Alcatel-Lucent. "MPLS Transport Profile Data Plane Architecture". RFC 5960 (Standards Track), 2010.

- [14] Fujitsu. Understanding pbb-te for carrier ethernet. [http :
//www.fujitsu.com/downloads/TEL/fnc/whitepapers/UnderstandingPBBTE.pdf](http://www.fujitsu.com/downloads/TEL/fnc/whitepapers/UnderstandingPBBTE.pdf), 2008.
- [15] Gilbert Held. *Carrier Ethernet Providing the Need for Speed*. 2008.
- [16] Abdul Kasim. *Delivering Carrier Ethernet Extending Ethernet beyond the LAN*. 2008.
- [17] Daniel Maier. *IP Multicast with Applications to IPTV and Mobile DVB-H*. Wiley-Interscience, New York, NY, USA, 2008.
- [18] L Martini, E. Rosen, N. El-Aawar, Cisco Systems, Tellabs, G. Heron, N. El-Aawar, and LLC Level 3 Communications. "Encapsulation Methods for Transport of Ethernet over MPLS Networks". RFC 4448 (Standards Track), 2006.
- [19] B. Niven-Jenkins. Requirements of an MPLS Transport Profile. RFC 5654 (Standards Track), 2010.
- [20] Gerard O'Driscoll. *Next Generation IPTV Services and Technologies*. Wiley-Interscience, New York, NY, USA, 2008.
- [21] Professor Armando Pinto. Apontamentos de redes Ópticas. [http:
elearning.ua.pt](http://elearning.ua.pt), 2008.
- [22] Lampros Raptis e Kostas Vaxevanakis Rafael Sanchez. Ethernet as a carrier grade technology: Developments and innovations. *IEEE Communications Magazine*, 22(1):88–94, 2008.
- [23] E. Rosen, Inc. Cisco Systems, A. Viswanathan, Inc. Force10 Networks, R. Callon, and Inc. Juniper Networks. "Multiprotocol Label Switching Architecture". RFC 3031 (Standards Track), 2001.
- [24] Kenneth H. Rosen. *Discrete Mathematics and Its Applications*. McGraw-Hill Higher Education, 2002.
- [25] Nishith Sinha and Ozm Ran. The statistics of switched broadcast. *SCTE 2005 Conference on Emerging Technologies*, 2005.
- [26] Francisco Vaz. *Probabilidades e Processos Estocásticos*. Universidade de Aveiro, Aveiro, Portugal, 2002.